Representable Options

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Matthias Lenga

Kiel, 2017

Referent:	**Prof. Dr. Jan Kallsen**
Koreferent:	**Prof. Dr. Josef Teichmann**

Tag der mündlichen Prüfung:	01.06.2017
Zum Druck genehmigt:	01.06.2017

gez. Prof. Dr. Natascha Oppelt, Dekanin

Acknowledgements

There is an old saying that "gratitude is not only the greatest of virtues, but the parent of all others." It occurs to me that in the process of pursuing the winding paths shaped by perpetually shifting ambitions, I have yet missed several opportunities to express my sincere appreciation to various people in my life. Finding the appropriate words which reflect my great esteem for their numerous good deeds seems to be a difficult, if not, an impossible task.

I am greatly indebted to my supervisor Prof. Dr. Jan Kallsen. His impressive knowledge, his unresting curiosity and his abiding commitment to teaching have always been a great inspiration to me. Jan Kallsen has always encouraged me to pursue my own ideas and what is more, he was very supportive of this and other endeavors during my time as doctoral student at the Faculty of Mathematics and Natural Sciences. In my view, it is no exaggeration to say that he is a true master of his art.

I want to thank my second referee, Prof. Dr. Josef Teichmann, for bringing the articles [JM1, JM2] to my attention and for giving me the opportunity to present my results at the Institute for Theoretical Studies of the ETH Zürich.

Moreover, I am grateful to Prof. Dr. Sören Christensen for taking interest in my research, for being an excellent teacher and for including me in several exciting side projects.

From the bottom of my heart, I would like to thank my parents, Kerstin and Rainer Lenga, for their love and for imparting to me the value of learning. Throughout my whole life I could always count on their untiring support and patient advice. Furthermore, I am grateful to my brother Stephan for the nice years we shared in Kiel and for our weekly discussions concerning computer science and all the other things that matter in life.

Abstract

We call a given American option representable if there exists a European claim which dominates the American payoff function at any time and such that the values of the two options coincide within the continuation set associated to the American option. This mathematical concept has interesting implications from a probabilistic, analytic, financial and numerical point of view. We aim at analyzing the notion of representability and linking it to embedded American payoffs in the sense of Jourdain and Martini and cheapest dominating European options originating from the work of Christensen. This process reveals a new duality structure between European and American valuation problems which we deem as very promising for future research. Relying on methods from convex optimization, we make a first step towards verifying representability of certain American claims. Furthermore, we discuss some computational aspects related to semi-infinite linear programming theory. This ultimately leads to an iterative procedure which generates upper and lower bounds for American option prices as well as a spline approximation to the early exercise boundary. The algorithm is benchmarked against high-precision methods from the literature.

Zusammenfassung

Wir bezeichnen ein amerikanisches Derivat als darstellbar, falls eine europäische Option existiert, welche innerhalb des Fortsetzungsgebietes preisgleich zu der amerikanischen Option ist und deren assoziierte Wertfunktion den amerikanischen Payoff zu jeder Zeit dominiert. Aus diesem Konzept lassen sich diverse Schlüsse ableiten, welche sowohl aus einer wahrscheinlichkeitstheoretischen oder finanzmathematischen Perspektive, als auch von einem analytischen oder numerischen Standpunkt aus betrachtet von Interesse sind. Die vorliegende Dissertation zielt darauf ab, mittels Darstellbarkeit eine Brücke zwischen eingebetteten amerikanischen Auszahlungen, im Sinne von Jourdain und Martini, und den von Christensen diskutierten billigst dominierenden europäischen Optionen, zu schlagen. Hierbei stoßen wir auf einen bisher unbekannten strukturellen Zusammenhang zwischen amerikanischer und europäischer Optionsbewertung. Diesen erachten wir als interessant und reichhaltig hinsichtlich zukünftiger Forschungsvorhaben. Für gewisse amerikanische Auszahlungsprofile wagen wir, unter Zuhilfenahme von Methoden der konvexen Optimierung, einen ersten Schritt in Richtung Lösung des Darstellbarkeitsproblems. Ergänzend diskutieren wir einige algorithmische Aspekte im Rahmen der Theorie semi-infiniter linearer Programme. Abschließend präsentieren wir ein iteratives Verfahren, welches sowohl obere und untere Schranken für amerikanische Optionspreise, als auch eine Spline Approximation der Ausübungsgrenze generiert. Zur Leistungsbemessung ziehen wir Präzisionsmethoden aus der Fachliteratur zu Rate.

Contents

Es irrt der Mensch, solang er strebt.
Faust I, Vers 317 / Der Herr

1 Introduction

1.1 Pricing of American options

It might be fair to say that *options* belong to the most important financial instruments in the world of modern finance. Basically speaking, an option is a contract which gives the holder the right to trade a certain number of assets according to some clearly specified terms. For example, European call and put options allow the holder to purchase or sell an underlying asset for a fixed price at a certain date. In general, options are complex securities with a versatile nature - some of them a great deal of risk, others constitute rather conservative investments. By incorporating these contracts into their portfolios, investors can flexibly shape the latter according to their individual needs. A lot of research in mathematical finance and the related fields was driven by the central question of how an economically reasonable price for such derivatives can be obtained. In 1973 Fischer Black and Myron Scholes rigorously derived in their seminal article [BS] explicit formulas for theoretical option prices in an idealized market model. More precisely, they consider a continuous-time model consisting of a riskless bond and a non-dividend paying risky stock. It is assumed that the bond price B and the asset price S evolve according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}B_t &= rB_t\mathrm{d}t, \\ \mathrm{d}S_t &= \mu S_t\mathrm{d}t + \sigma S_t\mathrm{d}W_t \end{aligned} \tag{1.1}$$

where $\mu, r \in \mathbb{R}, \sigma \in \mathbb{R}_{++}$ and W denotes a standard Brownian motion. Furthermore, they assume that at any time it is possible to buy and sell any amount of stock or bond without paying a transaction fee. Within this market setup, nowadays known as the celebrated Black-Scholes model, the just mentioned authors derive closed formulas for the prices and hedging positions associated to European put and call options.
In the same year Robert C. Merton published a follow-up article [ME] containing further ground-breaking insight on option pricing and important extensions of the Black-Scholes theory. In particular, he considers the valuation of American call and put options within the framework of [BS]. Unlike European derivates, American type contracts can be exercised by the holder at any time up to some prespecified expiration date. In [ME] it is shown that American call options will never be exercised prior to expiration. Basic arbitrage theory implies that the value of an American call must coincide with the option price of its European counterpart.
Under mild assumptions, which are satisfied by the market (1.1), the well-known call-put parity allows us to deduce the value of a European put option from the price of a European call and vice versa. Unfortunately, even in the very basic Black-Scholes

setting, a relation of this kind does not exist for American style put and call options. In [ME, Theorem 13] it is shown that, in contrast to call options, the price of an American put exceeds the value of its European counterpart due to premature exercising. For the perpetual American put, i.e. an American put option which never expires, Merton realized that the valuation problem reduces to solving an ordinary differential equation, and therefore an explicit representation of the option value can indeed be derived, cf. [ME, Section 8]. Despite tremendous efforts, it was impossible to obtain such explicit pricing formulas for American put options with finite maturities.

Leaving the Black-Scholes scenario does not necessarily improve the situation. Quite the contrary, for other continuous-time market models, analytic solutions of American or even European valuation problems are seldomly known. Consequently, the practitioner is in need of fast and reliable numerical tools. A brief glance into the bibliography of any standard textbook on numerical methods in financial mathematics, see for instance [SD], provides the reader with a first impression on the fast and extensive development of this research field. Summarizing the vast amount of literature within the scope of this thesis would be a futile endeavor. Nevertheless, we want to sketch three rudimentary numerical approaches towards American option pricing. The short survey below mainly focuses on the finite-maturity American put in the Black-Scholes setting and rather aims at providing some basic intuition while omitting any mathematical details. In this regard we ask the reader to be lenient with us. We write $g(s) := (K - s)_+$ for the put payoff with strike $K \in \mathbb{R}_{++}$. In the market model (1.1) the fair price of an American put on the risky asset is a function $v_{\mathrm{am},g}(\vartheta, s)$ depending only on the option's maturity $\vartheta \in \mathbb{R}_+$ and the spot price $s \in \mathbb{R}_{++}$ of the underlying, cf. [PS, Section 25]. If not explicitly stated otherwise, we denote by $T \in \mathbb{R}_{++}$ some finite terminal time.

- **Finite difference methods:** Exploiting the Markovian nature of (1.1), it can be shown that the value function $v_{\mathrm{am},g}$ associated to the American put satisfies the linear complementary problem (LCP)

$$\begin{aligned} &\mathcal{D} v_{\mathrm{am},g} \geq 0 && \text{a.e. in } (0,T] \times \mathbb{R}_{++}, \\ &v_{\mathrm{am},g} \geq g && \text{a.e. in } (0,T] \times \mathbb{R}_{++}, \\ &(v_{\mathrm{am},g} - g)\, \mathcal{D} v_{\mathrm{am},g} = 0 && \text{a.e. in } (0,T] \times \mathbb{R}_{++}, \\ &v_{\mathrm{am},g}(0,s) = g(s) \end{aligned} \tag{1.2}$$

  where $\mathcal{D} := \partial_\vartheta - rs\partial_s - \frac{1}{2}\sigma^2 s^2 \partial_{ss} + r$, see [LL, Theorem 5.3.4]. In order to approximate the value function, we first choose a suitable finite grid $(\vartheta_i, s_j) \in [0,T] \times \mathbb{R}_+$ where $i = 0, ..., M$ and $j = 0, ..., N$ with $\vartheta_0 = s_0 = 0$. Replacing all differential expressions in (1.2) by appropriate difference quotients and introducing suitable boundary conditions yields a sequence of discrete LCPs for the approximate values of $v_{\mathrm{am},g}$ at the grid points, cf. [SD, Subsection 4.6.1]. In case of the American put, the discrete LCPs appearing while iterating through the time layers can be solved very efficiently by the Brennan-Schwartz method. For differently structured payoff profiles, the discrete LCPs are frequently solved by iterative indirect methods, for instance the projected SOR algorithm, cf. [SD, Algorithm 4.11]. The Brennan-Schwartz algorithm originates from the 1977 article [BRS] which presents one of the earliest finite difference approximations to the American put value function. Even though the formulation of the LCP presented in [BRS] was erroneous, the proposed numerical method turned out to be justified. A thorough analysis of the Brennan-Schwartz approach and a general discussion of variational inequalities related to American option valuation can be found in [JLL]. Moreover, we refer the reader to the survey article [TV] on finite difference techniques in computational finance. As pointed out by [TV], the computational feasibility of such methods is strongly affected by the number of grid points. As a rule of thumb, the computational effort increases exponentially with the number of underlying assets. Consequently, finite difference based algorithms are deemed to be useful in low dimensional settings.

- **Tree approximations:** Tree methods rely on the construction of discrete Markov chains which approximate the stochastic dynamics of the underlying assets. The usage of tree models in finance originates from the work of Sharpe and the seminal article [CRR] by Cox, Ross and Rubinstein which contains fundamental insights on option pricing in binomial trees. The basic concept of binomial tree models is very intuitive: Choose some finite time grid $0 = t_0 < ... < t_N = T$ and suppose that the prices of a riskless bond B and some risky asset S move according to the equations
$$B_n = B_0 \exp(rt_n)$$
$$S_n = S_0 \prod_{l=1}^{n} (1 + Z_l)$$
for any $n \in \{0, ..., N\}$. Here we denote by $Z_1, ..., Z_N$ i.i.d. random variables satisfying $\mathbb{P}(Z_1 = u) = 1 - \mathbb{P}(Z_1 = d) = p$ with $p \in (0, 1)$ and $-1 < d < u$. Hence, the parameters u, d represent the possible one-period relative stock price changes and $r \in \mathbb{R}_{++}$ corresponds to a constant interest rate. Models of the latter type are well interpretable and mathematically simple. In particular, formulas for European and American option prices can be easily obtained, see [LL, Chapter 1 and 2]. We refer the reader to [SH] for a comprehensive didactic approach towards mathematical finance based on tree models. In [CRR] it is shown that we can choose parameter sequences $(r_k, u_k, d_k, p_k, N_k, (t_i^k)_{i=0,...,N_k})_{k\in\mathbb{N}}$ such that the associated sequence of binomial trees converges in distribution to the continuous-time Black-Scholes model. Under certain regularity assumptions, European and American option prices stemming from the discrete-time tree approximations converge to the corresponding prices in the limiting Black-Scholes market. We refer the reader to [KU] for a rather general treatment of discrete-time approximations to Markovian stochastic control problems. Short and easily comprehensible introductions to tree-based methods, as well as further references, can be found in [LZ] and [SD, Section 1.4]. A major drawback of tree approximations is that the computational

effort grows exponentially as the number of model variables increases. In practice, tree based algorithms are therefore predominantly useful for models with one or two underlying assets, cf. [LZ] and [BG].

- **Monte Carlo methods:** In addition to the finite difference approach put forward by Schwartz et al. and the CRR tree approximations, Phelim Boyle presented a third method for the numerical solution of option pricing problems in 1977. In order to obtain European prices, he proposed to evaluate the related integrals by Monte Carlo simulation. In addition, he recommended to apply certain variance reduction techniques in order to achieve a greater precision and smaller confidence intervals for a given sample size, cf. [BO]. For example: Calculating the price of some European option in the Black-Scholes model is equivalent to the evaluation of $\mathbb{E}[h(Z)]$ for some measurable function h and some normally distributed random variable Z. Under mild integrability assumptions, the law of large numbers yields that
$$\hat{S}_N := \frac{1}{N}\sum_{n=1}^{N} h(Z_n) \to \mathbb{E}\left[h(Z)\right]$$
as $N \to \infty$ where $Z_1, Z_2, ...$ denote i.i.d. random variables which follow the same distribution as Z. There exist several methods to increase the efficiency of Monte Carlo estimates. For instance, in [BO] the usage of antithetic variables or control variates is suggested. We refer the reader to [GL, Chapter 4] for a general overview of variance reduction methods.

  Unfortunately, Monte Carlo techniques for the pricing of American type options turn out to be more complex as their European counterparts. As a cornerstone of this topic we want to mention the random tree approach put forward by Broadie and Glasserman. In their article [BG] they first consider the valuation of *Bermudan* options in a Markovian setting, i.e. options which only allow for early exercise at some finite set of time points. In the following they argue that American option values may be extrapolated from the solutions of certain Bermudan type pricing problems. The main ingredients of their Bermudan pricing algorithm can be roughly summarized as follows:

  - First, a tree is simulated whose nodes correspond to simulations of the model state variables at the possible exercise times. By applying a dynamic programming algorithm to this tree, an estimator Θ_{H} for the Bermudan option value V is obtained. In [BG] it is shown that the estimator Θ_{H} is biased high, i.e. $\mathbb{E}\left[\Theta_{\mathrm{H}}\right] \geq V$. Moreover it is shown that Θ_{H} converges to V as we increase the number of branchings b at each node. Based on the same simulated tree, an estimator Θ_{L} is derived which is biased low and consistent in the sense above.
  - Averaging the high and low estimators generated from n independent simulations of trees with the same branching complexity b yields estimators $\overline{\Theta}_{\mathrm{H}}(n,b)$, $\overline{\Theta}_{\mathrm{L}}(n,b)$ and an asymptotic $1-\delta$ confidence interval for the value V of the

type

$$\left(\overline{\Theta}_{\mathrm{L}}(n,b) - \psi_1(\delta,n,b), \overline{\Theta}_{\mathrm{H}}(n,b) + \psi_2(\delta,n,b)\right)$$

where ψ_1, ψ_2 denote certain functions depending only on the parameters δ, n and b. The analysis of [BG] shows that the bounds of the latter interval can be arbitrarily tightened by increasing b and n.

The computational complexity of this method depends exponentially on the amount of possible exercise times. Consequently, the approach turns out to be numerically infeasible if this number is too large. We refer the reader to Section 8.3 of [GL] for a thorough discussion of this topic and some additional remarks concerning implementation and enhancements. Another method for Bermudan option pricing was proposed by Longstaff and Schwartz, cf. [LS]. Their algorithm relies on approximating the conditional expectation occurring in the dynamic programming equation associated to the valuation problem by a suitably chosen projection on some finite dimensional space of random variables. A detailed discussion of the Longstaff-Schwartz algorithm as well as a convergence analysis can be found in [CLP].

As Boyle already pointed out, Monte Carlo techniques turn out to be useful in many cases where finite difference methods or tree approximations fail, in particular in the presence of jumps or in high dimensional market models. In general, these methods do not generate global approximations of the value function $v_{\mathrm{am},g}$ associated to the pricing problem. Due to their high computational effort, they are therefore considered as predominantly useful when only a few option values are required. The textbook [GL] provides a very comprehensive survey of Monte Carlo methods and their financial applications.

Aside the numerous algorithmic aspects, the mathematical essence of American option pricing and the intimately connected problem of optimal stopping have been intensively studied during the last decades. It turned out to be fruitful to adopt two different mathematical viewpoints simultaneously. On the one hand, the problem can be formulated in the context of martingales, Snell envelopes and stopping times. On the other hand, by relying on Markov process theory, one can adopt more local perspective which allows to transcribe American valuation problems in terms of differential operators and free boundary problems.
We refer the reader to the monograph [PS] for an excellent treatment of optimal stopping theory and various applications. The exposition of G. Peskir and A. N. Shiryaev clearly distinguishes between the two different approaches and rigorously demonstrates the beneficial nature of their interplay. Moreover, the second chapter of the aforementioned textbook will serve us as the prime reference for all basic notions from stochastic process theory.

1.2 Contribution of this work

The thesis at hand aims at providing a new perspective on American type options which hopefully paves the way for new numerical methods. The central idea is to reduce the valuation of American options to the simpler problem of computing prices of European contingent claims whose payoffs do not depend on the paths of the underlying stochastic process.

For ease of exposition we consider the risk-neutral Black-Scholes setting consisting of a deterministic bond and some risk-bearing stock. The value of the bond B and the logarithmic asset price X evolves according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}B_t &= rB_t\mathrm{d}t, \\ \mathrm{d}X_t &= \left(r - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\mathrm{d}W_t, \end{aligned} \tag{1.3}$$

where $r \geq 0$, $\sigma > 0$, $B_0 := 1$ and $W = (W_t)_{t\in\mathbb{R}_+}$ denotes a standard Brownian motion. The filtration generated by the process W is denoted by $\mathcal{F} := (\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Following the notational conventions from Markov process theory, we denote by $(\mathbb{P}_x)_{x\in\mathbb{R}}$ the family of measures such that $\mathbb{P}_x(X_0 = x) = 1$ holds true for any $x \in \mathbb{R}$. We write $\mathcal{T}$ for the set of stopping times satisfying $\mathbb{P}_x(0 \leq \tau < \infty) = 1$ for all $x \in \mathbb{R}$. Moreover, for any $\vartheta \in \mathbb{R}_+$ the aggregate of all $[0,\vartheta]$-valued stopping times is denoted by $\mathcal{T}_{[0,\vartheta]}$. Suppose that $f : \mathbb{R} \to \mathbb{R}_+$ is a measurable payoff function. The fair *value of a European option* with payoff $f(X)$, maturity $\vartheta \in \mathbb{R}_+$ and initial logarithmic stock price $x \in \mathbb{R}$ is denoted by $v_{\mathrm{eu},f}(\vartheta, x)$, i.e.

$$v_{\mathrm{eu},f}(\vartheta, x) := \mathbb{E}_x\left[e^{-r\vartheta} f(X_\vartheta)\right]. \tag{1.4}$$

Similarly, for a continuous payoff function $g : \mathbb{R} \to \mathbb{R}_+$ satisfying the integrability condition

$$\mathbb{E}_x\left[\sup_{t\in[0,T]} e^{-rt} g(X_t)\right] < \infty, \tag{1.5}$$

for some time horizon $T \in [0,\infty]$, the fair *value of an American claim* with payoff process $Z = g(X)$, maturity $\vartheta \in [0,T]$ and logarithmic spot price $x \in \mathbb{R}$ is denoted by $v_{\mathrm{am},g}(\vartheta, x)$, i.e.

$$v_{\mathrm{am},g}(\vartheta, x) := \sup_{\tau\in\mathcal{T}} \mathbb{E}_x\left[e^{-r(\tau\wedge\vartheta)} g(X_{\tau\wedge\vartheta})\right]. \tag{1.6}$$

We will call $v_{\mathrm{eu},f}$ the *European value function* associated to f and $v_{\mathrm{am},g}$ the *American value function* associated to g. Until the end of this section let us assume that T is finite. If not explicitly stated otherwise, the value functions from (1.4) and (1.6) will always be parametrized in maturity/log-price coordinates. Condition (1.5) warrants that $v_{\mathrm{am},g}$ is finitely valued and lower semi-continuous on the set $[0,T] \times \mathbb{R}$, cf. Section 5.3. Indeed, for any stopping time $\tau \in \mathcal{T}$ we find by dominated convergence that the mapping $(\vartheta, x) \mapsto \mathbb{E}_x\left[e^{-r(\tau\wedge\vartheta)} g(X_{\tau\wedge\vartheta})\right]$ is continuous on $[0,T] \times \mathbb{R}$. Lemma 5.9 now directly implies the lower semi-continuity of the American value function. Following [PS], we write

$$C_T := \{(\vartheta, x) \in \mathbb{R}_+ \times \mathbb{R} \mid \vartheta \leq T \text{ and } v_{\mathrm{am},g}(\vartheta, x) > g(x)\} \tag{1.7}$$

for the *continuation region* and

$$S_T := C_T^{\mathrm{c}} = \{(\vartheta, x) \in \mathbb{R}_+ \times \mathbb{R} \mid \vartheta \leq T \text{ and } v_{\mathrm{am},g}(\vartheta, x) = g(x)\} \tag{1.8}$$

for the *stopping region* associated to the American claim. We explicitly remark that the sets C_T and S_T will always be parametrized by maturity/log-price coordinates. In case that the time horizon T is finite, one can apply the transformation $\vartheta \mapsto T - \vartheta$ in order to switch to calender time.

Now, fix some terminal time $T \in \mathbb{R}_{++}$ and some log-price $x_0 \in \mathbb{R}$ such that the point (T, x_0) is contained in C_T. For this introductory section let us assume that C_T is a connected set. We say that a European payoff function $f : \mathbb{R} \to \mathbb{R}_+$ *represents* the American payoff function $g : \mathbb{R} \to \mathbb{R}_+$ if the European value function associated to f dominates the value of the American option everywhere and these two functions coincide within the continuation region of the American claim, i.e. we have $v_{\mathrm{eu},f} \geq v_{\mathrm{am},g}$ on the set $[0, T] \times \mathbb{R}$ and $v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x)$ holds true for any $(\vartheta, x) \in C_T$. The following guiding question for the thesis at hand arises:

Given an American payoff function g, is there a European payoff function f representing g ?

If true, we call the payoff g *representable* and f the representing European claim. The concept of representability has several interesting implications from a probabilistic, analytic and financial point of view, for example:

- Global approximations of the American value function can be efficiently computed by means of linear programming, cf. Chapter 3.
- A buy-and-hold position in the European option with time T payoff $f(X_T)$ hedges the American claim perfectly.
- Within the continuation region C_T, the difference $v_{\mathrm{am},g} - v_{\mathrm{eu},g}$ corresponds to the fair value of a European payoff with time T payoff $h(X_T) := f(X_T) - g(X_T)$, i.e.

  $$v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},g}(\vartheta, x) + v_{\mathrm{eu},h}(\vartheta, x) \tag{1.9}$$

  for any $(\vartheta, x) \in C_T$. To put differently, the early exercise premium of the American option can be interpreted as the price of a European claim with a specific payoff profile.

- The Snell envelope corresponding to the American option allows for a Markovian-style decomposition, see Equation (1.16) below.
- Certain analytical properties of the European value function associated to the representing payoff f transfer to the American value function $v_{\mathrm{am},g}$, cf. Subsection 2.1.1 and Section 2.2.
- Some analytical properties of the early exercise curve can be obtained easily. Indeed, the latter coincides with the boundary of the set
$$\{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid v_{\mathrm{eu},f}(\vartheta, x) = g(x)\}.$$
This allows to derive smoothness properties of the early exercise curve from the analyticity of $v_{\mathrm{eu},f}$ by means of the implicit function theorem, cf. Section 2.2.
- The solution of the free boundary problem associated to the American option can be extended to a solution of the Black-Scholes partial differential equation beyond the free boundary.

On top of verifying the representability of a given option, one may ask how to obtain the representing European payoff, at least numerically. Moreover, are possibly all American options representable? Or, if this is not the case, do representable options exist at all - except for the obvious case where early exercise is suboptimal and hence g itself represents g?

The concept of representability is not studied here for the first time. It was considered in two articles by Jourdain and Martini, which have not yet received the attention they deserve. In [JM1] it is shown that many European payoffs represent some American claim which can be obtained in a natural way. Indeed, given some European payoff function f, they define an American payoff function $\mathrm{am}_T(f) : \mathbb{R} \to \mathbb{R}_+$ as

$$\mathrm{am}_T(f)(x) := \inf_{\vartheta \in [0,T]} v_{\mathrm{eu},f}(\vartheta, x), \tag{1.10}$$

from now on called the *embedded American option* (EAO) associated to f. If the infimum in (1.10) is attained in a connected curve, then f represents its embedded American option $\mathrm{am}_T(f)$, cf. [JM1, Theorem 5]. Jourdain and Martini provide an explicit example where this is the case. Additionally, they show that embedded American payoff functions satisfy certain analyticity properties, cf. [JM1, Proposition 16]. From their results we conclude that representable options exist but that not all American payoff functions are representable. In their follow-up article [JM2] the aforementioned authors study the American put option in detail. They show that it cannot be represented by any of a seemingly general and reasonable candidate family of European contingent claims. This may be considered as an indication that this particular option may not be representable after all. Summing up, the main contribution of Jourdain and Martini is to provide a way to obtain an American payoff function g that is represented by a given European claim f.

Our question here is rather the converse: Given g, is there a representing European claim f, and how can it be obtained? In order to tackle these problems, we make use of a concept originating from [CR]. The key contribution of [CR] is the linear optimization problem

$$\begin{aligned} &\text{minimize} && v_{\mathrm{eu},f}(T, x_0) \\ &\text{subject to} && f : \mathbb{R} \to \mathbb{R}_+ \text{ measurable and} \\ & && v_{\mathrm{eu},f}(\vartheta, x) \geq g(x) \text{ for all } (\vartheta, x) \in [0,T] \times \mathbb{R}. \end{aligned} \tag{1.11}$$

We call a minimizer f of (1.11) *cheapest dominating European option* (CDEO) of g relative to (T, x_0). The infinite dimensional linear problem (1.11) can be numerically solved by semi-infinite programming methods, cf. [HK], [IW] and Chapter 3 below. It is easy to see that the fair price of a CDEO f provides an upper bound to the value of the given American claim g. However, in [CR] it remains open how large the gap between the two actually is. While there is a priori no reason why the two should coincide, numerical studies in [CR] indicate that for certain American payoffs the difference seems to be very small. In the present thesis we use the CDEO associated to some American option as the natural candidate for the generating European payoff. To put differently, constructing the CDEO constitutes in some sense the inverse relation to the embedding operation, cf. Proposition 2.10 and Theorem 2.52. Indeed, if g is representable at all, it must be represented by its CDEO, as we will see below. This also answers the question how to obtain a representing European payoff function numerically if it exists at all, cf. Section 3.1.

It is important to distinguish the minimization problem (1.11) and more generally the present study from the well-known duality approaches put forward by [RG], [DK] and [KH]. Consider again an American payoff function $g : \mathbb{R} \to \mathbb{R}_+$ leading to the discounted exercise process $\hat{Z}_t := e^{-rt} g(X_t)$ and some finite time horizon $T \in \mathbb{R}_{++}$. From [RG, Theorem 1] we know that

$$v_{\mathrm{am},g}(T, x_0) = \inf \left\{ \mathbb{E}_{x_0} \left[\sup_{t \in [0,T]} (\hat{Z}_t - M_t) \right] \;\middle|\; M \text{ martingale with } M_0 = 0 \right\}. \tag{1.12}$$

Indeed, the inequality $\leq$ is obvious because

$$\mathbb{E}_{x_0}[\hat{Z}_\tau] = \mathbb{E}_{x_0}[\hat{Z}_\tau - M_\tau] \leq \mathbb{E}_{x_0} \left[\sup_{t \in [0,T]} (\hat{Z}_t - M_t) \right]$$

for any $[0, T]$-valued stopping time τ and any martingale M with $M_0 = 0$. For the converse inequality consider the Doob-Meyer decomposition

$$V = V_0 + M^V - A^V \tag{1.13}$$

of the Snell envelope V of the discounted exercise process $\hat{Z}$, i.e. M^V is a martingale and A^V an increasing process with $A_0^V = M_0^V = 0$, cf. [PS, Theorem 3.1]. It is well-known that V corresponds to the discounted fair price process associated to the American claim, i.e. $V_t = e^{-rt} v_{\mathrm{am},g}(T-t, X_t)$, cf. [LB, Theorem 4.1.1]. Since

$$\hat{Z}_t - M_t^V \leq V_t - M_t^V = V_0 - A_t^V \leq V_0 = v_{\mathrm{am},g}(T, x_0)$$

for any $t \in [0,T]$, we conclude that the inequality $\geq$ holds in (1.12) as well. Similarly, observe that

$$v_{\mathrm{am},g}(T, x_0) = \inf \left\{ \mathbb{E}_{x_0}[Y] \;\middle|\; \begin{array}{l} Y \geq 0 \text{ random variable with} \\ \hat{Z}_t \leq \mathbb{E}_{x_0}[Y|\mathcal{F}_t] \ \forall t \in [0,T] \end{array} \right\}. \tag{1.14}$$

Again, the inequality $\leq$ is obvious because any martingale dominating $\hat{Z}$ majorizes the discounted American option price process V. The converse inequality $\geq$ follows from choosing $Y = V_0 + M_T^V$, where V and M^V are defined as above. The linear problem (1.11) can be rephrased as

$$\inf \left\{ \mathbb{E}_{x_0}\left[e^{-rT} f(X_T)\right] \;\middle|\; \begin{array}{l} f : \mathbb{R} \to \mathbb{R}_+ \text{ measurable with} \\ \hat{Z}_t \leq \mathbb{E}_{x_0}\left[e^{-rT} f(X_T) \middle| \mathcal{F}_t\right] \ \forall t \in [0,T] \end{array} \right\}, \tag{1.15}$$

which seems almost identical to the right-hand side of (1.14). However, the dominating European payoff Y in (1.14) may well be path dependent, which is not the case in (1.15). And indeed, it is easy to see that the terminal value $V_0 + M_T^V$ cannot typically be written as a function of X_T, e.g. in the case of an American put. Therefore, the identities (1.12) and (1.14) do not help in deciding whether the value of the CDEO in the sense of (1.11) coincides with the price of the American option at hand.
From a different perspective, one may note that the martingale in the Doob-Meyer decomposition (1.13) is not the only one that leads to optimal choices in (1.12) and (1.14). In fact, we could replace M^V by $\widetilde{M}$ in any decomposition of the form

$$V = V_0 + \widetilde{M} - \widetilde{A} \tag{1.16}$$

with some martingale $\widetilde{M}$ and some non-negative process $\widetilde{A}$ satisfying $\widetilde{A}_0 = \widetilde{M}_0 = 0$. Contrary to the unique decomposition (1.13), we do not require $\widetilde{A}$ to be increasing. As noted above, (1.15) coincides with the American option price (1.14) if we can choose $\widetilde{M}$ such that $V_0 + \widetilde{M}_T = e^{-rT} f(X_T)$ for some deterministic function f. In this case, the decomposition (1.16) is *Markovian-style* in the sense that both $\widetilde{M}_t$ and $\widetilde{A}_t$ are functions of t and X_t, i.e.

$$\begin{aligned} \widetilde{M}_t &= e^{-rT} \mathbb{E}_{X_t}\left[f(X_{T-t})\right] - V_0 \\ \widetilde{A}_t &= e^{-rT} \mathbb{E}_{X_t}\left[f(X_{T-t})\right] - e^{-rt} v_{\mathrm{am},g}(T-t, X_t) \end{aligned}$$

at any time t. Hence, the issue of representability is linked to the existence of Markovian-style decompositions (1.16) of the Snell envelope associated to the optimal stopping problem.

The present study serves different purposes. In Section 2.1 we establish the link between embedded American options from [JM1], cheapest dominating European options from [CR] and the notion of representability. Subsection 2.1.2 contains several examples of representable American claims. In particular, we will provide an example of an embedded American payoff $g = \mathrm{am}_T(f)$ which is not represented by the generating European claim f. Furthermore, by providing an example, we show that representability may depend on the time horizon of the market, cf. Subsection 2.1.3. The main contribution of the thesis is contained in Section 2.2. First, we establish the existence of CDEOs in a distributional sense for sufficiently regular American payoff functions g. Secondly and more importantly, we provide a sufficient criterion which warrants that a given American claim is representable. The assumptions of this verification theorem depend on qualitative properties of the corresponding CDEO. Our numerical experiments indicate that these indeed seem to be satisfied for the American put in the Black-Scholes market, cf. Section 3.2. Somewhat independent of the question concerning the representability of American options, Section 3.1 outlines how CDEOs can be obtained numerically. Building on the well-established theory of semi-infinite programming, we approximate the infinite dimensional optimization task (1.15) by finite dimensional linear programs which can be solved with standard methods. Moreover, we provide supplementary convergence and consistency results. Afterwards, the precision and computational effort of the CDEO algorithm is benchmarked against high-precision methods found in the literature, cf. Section 3.3. As complement to the CDEO upper bound of the American option price, we will present a new algorithm which generates lower price bounds. The key idea is to generate an approximation to the early exercise boundary based on the dual optimizer associated to the CDEO, cf. Section 3.4.

1.3 Notation

Given some metric space S, we write $\mathcal{B}(S)$ for the Borel σ-algebra on S and $\mathcal{M}(S)$ for the vector space of regular Borel measures with finite total variation. The total variation of a measure μ will be denoted by $\|\mu\|$. For any $x \in S$ we denote by δ_x the Dirac unit mass concentrated at the point x. If R is another metric space, we write $\mathcal{B}(R, S)$ for the vector space of $\mathcal{B}(R)$-$\mathcal{B}(S)$-measurable mappings from R to S. The vector spaces of real-valued continuous functions, bounded continuous functions, continuous functions vanishing at infinity and compactly supported continuous functions on S are denoted by $C(S), C_b(S), C_0(S)$ and $C_c(S)$, respectively. The latter three are Banach spaces with respect to the norm $\|f\|_\infty := \sup_{x\in S} |f(x)|$ which generates the topology of uniform convergence $\mathcal{T}_{\mathrm{uc}}$. Furthermore, we denote by $\mathcal{M}^+(S), C^+(S), C_b^+(S), C_0^+(S)$ and $C_c^+(S)$ the cones of non-negative elements in the corresponding spaces. In a similar fashion we write $\mathbb{R}_+ := [0,\infty)$ and $\mathbb{R}_{++} := (0,\infty)$. Given two Banach spaces X and Y, we denote by $C^k(X,Y)$ the k-times continuously Fréchet differentiable mappings from X to Y. For a normed vector space $(V, \|\cdot\|)$ and any $x \in V, r \in \mathbb{R}_{++}$ we define the closed balls $B_V(x,r) := \{v \in V \mid \|v - x\| \leq r\}$, $B_V := B_V(0,1)$ and $B_V(r) := B_V(0,r)$. For $p \in [1,\infty)$ and a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$ we denote by $\mathrm{L}_p(\Omega, \mathcal{F}, \mu)$ the Banach space of p-integrable real-valued functions with respect to μ. The associated p-norm is denoted by $\|h\|_{p,\mu} := (\int_\Omega |h|^p \mathrm{d}\mu)^{1/p}$. The p-norm associated to the Lebesgue measure on $\mathbb{R}^n$ is simply denoted by $\|\cdot\|_p$. The Banach space of μ-essentially bounded functions is denoted by $\mathrm{L}_\infty(\Omega, \mathcal{F}, \mu)$. Sometimes the underlying space Ω and the sigma algebra $\mathcal{F}$ will be omitted in this notation. For an arbitrary set A, we define the following indicator functions:

$$\mathbb{1}_A(x) := \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \notin A. \end{cases} \qquad \mathcal{I}_A(x) := \begin{cases} 0 & \text{if } x \in A, \\ \infty & \text{if } x \notin A. \end{cases}$$

A normally distributed random variable with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}_{++}$ is denoted by $\mathcal{N}(\mu, \sigma^2)$. We write $\mathrm{N}(\mu, \sigma^2, \cdot)$ for the probability density function of $\mathcal{N}(\mu, \sigma^2)$. Moreover, we write $\varphi := \mathrm{N}(0, 1, \cdot)$ for the probability density function and Φ for the cumulative distribution function of a standard normal random variable. Given some Borel set $B \subset \mathbb{R}^d$, we denote by $\mathcal{U}_B$ a random variable which is uniformly distributed on B. The law of any random variable X is denoted by $\mathcal{L}(X)$. The value functions of a European and an American claim with payoff ϕ are denoted by $v_{\mathrm{eu},\phi}$ and $v_{\mathrm{am},\phi}$ respectively, see the definitions from Section 1.2. The closure and the interior of a set M in some topological space are denoted by $\operatorname{cl} M$ and $\operatorname{int} M$. We write $\partial M := \operatorname{cl} M \setminus \operatorname{int} M$ for the topological boundary of M. For any point x in some topological space, we denote by $\mathcal{U}(x)$ the system of all open sets containing x. The Euclidean topology on $\mathbb{R}^n$ is denoted by $\mathcal{T}_{\mathbb{R}^n}$. Furthermore, we agree upon the convention that $\inf \emptyset = +\infty$ and $\sup \emptyset = -\infty$.

2 A new duality between European and American options

2.1 Representable options

The objective of this section is to present some general facts about embedded, cheapest dominating, and representable options. In particular, we aim at providing some basic insight on the interplay of these mathematical notions.

2.1.1 Embedded American and cheapest dominating European options

Once again we consider the univariate Black-Scholes market (1.3). For European payoffs f and upper semi-continuous American payoffs g, we use the notation (1.4, 1.6) from Section 1.2 for the associated value functions. Note that $v_{\mathrm{am},g}(\vartheta, x)$ only deserves to be called fair option price if the integrability condition (1.5) holds.

2.1 Definition: *Let $f : \mathbb{R} \to \mathbb{R}_+$ denote a measurable European payoff and $g : \mathbb{R} \to \mathbb{R}_+$ an upper semi-continuous American payoff satisfying* (1.5)*.*

1. *The* embedded American option (EAO) *up to time $T \in [0, \infty]$ associated to f is defined as the payoff function* $\mathrm{am}_T(f) : \mathbb{R} \to \mathbb{R}_+$ *given by*

$$\mathrm{am}_T(f)(x) := \inf \{ v_{\mathrm{eu},f}(\vartheta, x) \mid \vartheta \in [0, T] \text{ and } \vartheta < \infty \} . \tag{2.2}$$

2. *We say that f* superreplicates *g up to $T \in [0, \infty]$ if the inequality $v_{\mathrm{eu},f}(\vartheta, x) \geq g(x)$ holds for any finite $\vartheta \in [0, T]$ and any $x \in \mathbb{R}$.*

3. *Given an initial logarithmic stock price $X_0 = x_0 \in \mathbb{R}$ and some finite time horizon $T \in \mathbb{R}_{++}$, we call a European payoff function $f^\star$* cheapest dominating European option (CDEO) *of g relative to T, x_0 if $f^\star$ superreplicates g up to T and $v_{\mathrm{eu},f^\star}(T, x_0) \leq v_{\mathrm{eu},f}(T, x_0)$ holds for any other European payoff f superreplicating g up to T.*

   Moreover, the set of all such CDEOs is denoted by $\mathrm{eu}_{T,x_0}(g)$*. We write* $\mathrm{eu}_{T,x_0}(g) = f^\star$ *if there is a unique CDEO $f^\star$, i.e. if* $\mathrm{eu}_{T,x_0}(g) = \{f^\star\}$*. Here we identify functions which only differ on a set of zero Lebesgue measure.*

Unless explicitly stated otherwise, we will consider finite time horizons $T \in \mathbb{R}_{++}$. Let us derive some direct consequences from the preceding definition.

2.3 Proposition: *Fix $(T, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$ and let $f, \tilde{f} : \mathbb{R} \to \mathbb{R}_+$ denote measurable European payoffs. Furthermore, suppose that $g : \mathbb{R} \to \mathbb{R}_+$ is an upper semi-continuous American payoff function satisfying the integrability condition* (1.5)*. Then:*

1. *The set $\mathrm{eu}_{T,x_0}(g)$ is convex.*
2. *If f superreplicates g up to time T, we have*
$$g(x) \leq v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{eu},f}(\vartheta, x) \tag{2.4}$$
for any $(\vartheta, x) \in [0, T] \times \mathbb{R}$ and consequently $g \leq \mathrm{am}_T(f) \leq f$. In particular, we have $g \leq \mathrm{am}_T(h)$ for any European payoff $h \in \mathrm{eu}_{T,x_0}(g)$.
3. *The mapping $T \mapsto \mathrm{am}_T(f)$ is decreasing and $f \mapsto \mathrm{am}_T(f)$ is increasing with respect to the natural ordering on $\mathbb{R}_+^{\mathbb{R}}$.*
4. *If $x \mapsto v_{\mathrm{eu},f}(\vartheta, x)$ constitutes for any $\vartheta \in [0, T]$ an upper semi-continuous mapping, the function $x \mapsto \mathrm{am}_T(f)(x)$ is upper semi-continuous as well.*
5. *For any $\lambda, \tilde{\lambda} \in \mathbb{R}_+$ we have $\mathrm{am}_T(\lambda f + \tilde{\lambda}\tilde{f}) \geq \lambda\, \mathrm{am}_T(f) + \tilde{\lambda}\, \mathrm{am}_T(\tilde{f})$.*
6. *For arbitrary $x \in \mathbb{R}$ we have*
$$|\mathrm{am}_T(f)(x) - \mathrm{am}_T(\tilde{f})(x)| \leq \sup_{\vartheta \in [0,T]} |v_{\mathrm{eu},f-\tilde{f}}(\vartheta, x)|. \tag{2.5}$$
Moreover, for any $p \in (1, \infty]$ the strong estimate
$$\|\mathrm{am}_T(f) - \mathrm{am}_T(\tilde{f})\|_p \leq c_p\, \|f - \tilde{f}\|_p \tag{2.6}$$
holds true for some constant c_p not depending on the functions f and $\tilde{f}$. In particular, we have $c_\infty = 1$.

Proof.

1. Choose $f_1, f_2 \in \mathrm{eu}_{T,x_0}(g)$ and note that for any $\lambda \in (0,1)$ the convex combination $f_\lambda := \lambda f_1 + (1-\lambda) f_2$ superreplicates g up to T. Moreover, we have $v_{\mathrm{eu},f_\lambda}(T, x_0) = \lambda v_{\mathrm{eu},f_1}(T, x_0) + (1-\lambda) v_{\mathrm{eu},f_2}(T, x_0) = v_{\mathrm{eu},f_1}(T, x_0)$ which implies that the payoff f_λ is indeed contained in $\mathrm{eu}_{T,x_0}(g)$.
2. Recall that for any $\vartheta \in (0, T]$ the discounted European value process
$$\left(e^{-rt} v_{\mathrm{eu},f}(\vartheta - t, X_t)\right)_{t \in [0,\vartheta]}$$
is a martingale on the time segment $[0, \vartheta]$. Indeed, applying the Markov property yields $e^{-rt} v_{\mathrm{eu},f}(\vartheta - t, X_t) = e^{-\vartheta r}\mathbb{E}_{X_t}\left[f(X_{\vartheta - t})\right] = e^{-\vartheta r}\mathbb{E}_x\left[f(X_\vartheta)|\mathcal{F}_t\right]$ for any $t \in [0, \vartheta]$. Owing to the superreplication property and the optional sampling theorem, we find that
$$v_{\mathrm{am},g}(\vartheta, x) = \sup_{\tau \in \mathcal{T}_{[0,\vartheta]}} \mathbb{E}_x\left[e^{-r\tau} g(X_\tau)\right]$$

$$\leq \sup_{\tau \in \mathcal{T}_{[0,\vartheta]}} \mathbb{E}_x \left[e^{-r\tau} v_{\mathrm{eu},f}(\vartheta - \tau, X_\tau) \right]$$
$$= v_{\mathrm{eu},f}(\vartheta, x)$$

holds true for any $(\vartheta, x) \in [0,T] \times \mathbb{R}$, which proves (2.4). Minimizing both sides of the latter inequality with respect to $\vartheta \in [0,T]$ yields $g(x) \leq \mathrm{am}_T(f)(x) \leq f(x)$ for any $x \in \mathbb{R}$.

3. This is obvious.

4. The point-wise infimum of an upper semi-continuous function collection constitutes an upper semi-continuous mapping, cf. Lemma 5.9.

5. This is obvious.

6. For the proof of this assertion we note that

$$\begin{aligned}
\mathrm{am}_T(f)(x) - \mathrm{am}_T(\tilde{f})(x) &= \sup_{\tilde{\vartheta} \in [0,T]} \inf_{\vartheta \in [0,T]} \Big(v_{\mathrm{eu},f}(\vartheta, x) - v_{\mathrm{eu},\tilde{f}}(\tilde{\vartheta}, x) \Big) \\
&\leq \sup_{\tilde{\vartheta} \in [0,T]} v_{\mathrm{eu},f-\tilde{f}}(\tilde{\vartheta}, x) \\
&\leq \sup_{\tilde{\vartheta} \in [0,T]} |v_{\mathrm{eu},f-\tilde{f}}(\tilde{\vartheta}, x)|.
\end{aligned}$$

Interchanging the roles of f and $\tilde{f}$ yields the point-wise estimate (2.5). The latter implies

$$\|\mathrm{am}_T(f) - \mathrm{am}_T(\tilde{f})\|_\infty \leq \sup_{x \in \mathbb{R}} \sup_{\vartheta \in [0,T]} v_{\mathrm{eu},|f-\tilde{f}|}(\vartheta, x) \leq \|f - \tilde{f}\|_\infty$$

which proves the strong estimate (2.6) for $p = \infty$. Now suppose that $p \in (1, \infty)$. Young's inequality shows that the linear operators

$$Q_\vartheta : \mathrm{L}_p \to \mathrm{L}_p \ ; \ h \mapsto v_{\mathrm{eu},h}(\vartheta, \cdot\,)$$

define a strongly continuous contraction semi-group, cf. [LA, p. 224, Theorem 1.2]. Lemma 5.3 allows us to conclude that the semi-group $(Q_\vartheta)_{\vartheta \in \mathbb{R}_+}$ is bounded analytic of angle $\frac{\pi}{2}$ in the sense of [EN, Definition 4.5]. Alternatively, the analyticity of Q_ϑ can be obtained from [EN, Corollary 4.9] by noting that the generator of $e^{r\vartheta} Q_\vartheta$ coincides on $C^2 \cap \mathrm{L}_p$ with the square of the generator associated to a certain strongly continuous group. In virtue of the point-wise estimate (2.5) we find that

$$\|\mathrm{am}_T(f) - \mathrm{am}_T(\tilde{f})\|_p \leq \left\| \sup_{\vartheta \in [0,T]} Q_\vartheta |f - \tilde{f}| \right\|_p .$$

If the semi-group Q_ϑ is self-adjoint on L_2, i.e. $r = \frac{\sigma^2}{2}$, the classic maximal theorem of E. M. Stein, cf. [ST, Chapter 3], implies that there exists some positive constant c_p which does not depend on the payoff functions $f, \tilde{f}$ such that

$$\left\| \sup_{\vartheta \in [0,T]} Q_\vartheta |f - \tilde{f}| \right\|_p \leq c_p \|f - \tilde{f}\|_p .$$

In case that the operators Q_ϑ are not self-adjoint on L_2, we can apply [MX, Corollary 4.2] in order to obtain an analogous estimate.

□

Let us remark that the validity of the latter proposition does not depend on the specific distributional properties of the Black-Scholes model. This is obvious for the Assertions 1 to 5, the weak estimate (2.5) and the strong estimate (2.6) for $p = \infty$. In case that $p \in (1, \infty)$, the validity of (2.6) essentially depends on the analyticity and the L_p-contractivity of the pricing semi-group Q_ϑ. Aside from the Lebesgue measure, the reader may think of other measures $\mu \in \mathcal{M}^+(\mathbb{R})$ such that [MX, Corollary 4.2] is applicable in order to obtain estimates of the type (2.6) for different p-norms $\|h\|_{p,\mu} := (\int_\mathbb{R} |h|^p \mathrm{d}\mu)^{1/p}$.

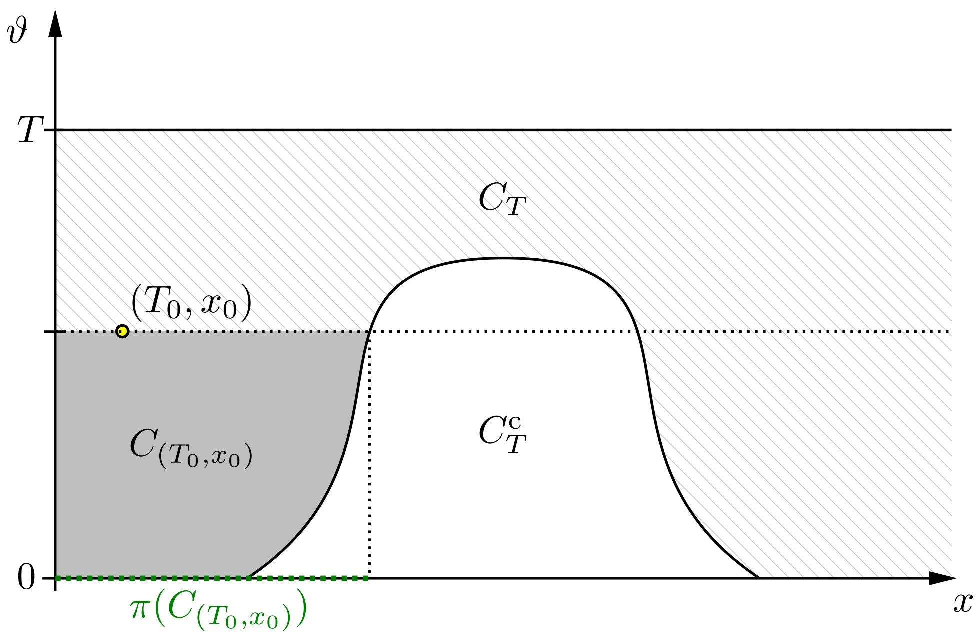

Figure 2.1: The sets $C_T, C_T^\mathrm{c}, C_{(T_0,x_0)}$ and $\pi(C_{(T_0,x_0)})$.

Now we turn to the *representability* of an American claim as explained in Section 1.2. To this end, we fix a terminal time $T \in \mathbb{R}_{++}$ and some continuous American payoff function $g : \mathbb{R} \to \mathbb{R}_+$ satisfying $g(x) \leq C(1 + |x|^k)$ for some constants $C, k \in \mathbb{R}_{++}$. Clearly, the payoff g satisfies the integrability condition (1.5) and [LB, Theorem 4.1.1] warrants that the associated value function $v_{\mathrm{am},g}$ is continuous. Moreover, we assume that the associated continuation region C_T, as defined in (1.7), is not empty and we denote by $S_T := C_T^\mathrm{c} = ([0,T] \times \mathbb{R}) \setminus C_T$ the corresponding stopping region. The set S_T is closed and for any maturity $\vartheta \in [0,T]$ the stopping time

$$\tau_\vartheta := \inf \{t \geq 0 \mid v_{\mathrm{am},g}(\vartheta - t, X_t) = g(X_t)\} \wedge \vartheta \tag{2.7}$$

is optimal for the stopping problem inherent in (1.6), cf. [PS, Corollary 2.9]. Given any $(T_0, x_0) \in C_T$ we denote by $C_{(T_0,x_0)}$ the connected component of the set $C_{T_0} = C_T \cap ([0, T_0] \times \mathbb{R})$ containing the point (T_0, x_0), see Figure 2.1.

2.8 Definition:

1. *We say that a European payoff f* (locally) represents *g relative to some point $(T_0, x_0) \in C_T$ if the function f superreplicates g up to time T_0 and $v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x)$ holds true for any $(\vartheta, x) \in C_{(T_0,x_0)}$. In this case we write*

$$f \overset{(T_0,x_0)}{\longrightarrow} g$$

and call g (locally) representable *relative to (T_0, x_0).*

2. *If $f \overset{(T_0,x_0)}{\longrightarrow} g$ for every $(T_0, x_0) \in C_T$, we say that f* (globally) represents *g up to time T. In this case we write*

$$f \overset{T}{\longrightarrow} g$$

and call g (globally) representable *up to time T.*

Let us emphasize that generally speaking, representability depends in a *local* manner on the connected components of the continuation region. That is to say, in Example 2.25 we will construct an American payoff which is representable with respect to each connected component of the associated continuation set, but which is not globally representable by any European claim. Furthermore, in Subsection 2.1.3 we will see that the representability of an American payoff may depend on the terminal time of the model. More precisely, we will show that the EAO associated to the European put in the Black-Scholes market is representable up to some maximal time horizon.
The following proposition collects some basic conclusions from the concept of representability and establishes a link between EAOs and CDEOs. A generalization of the second and third assertion for *measure type European claims* can be found in Lemma 3.2. For any set $M \subset \mathbb{R}_+ \times \mathbb{R}$ let us denote by

$$\pi(M) := \{x \in \mathbb{R} \mid (\vartheta, x) \in M \text{ for some } \vartheta \in \mathbb{R}_+\} \tag{2.9}$$

the projection of M onto the second coordinate, see for example Figure 2.1. Henceforth, functions which coincide up to a Lebesgue nullset will be implicitly identified whenever necessary.

2.10 Proposition: *Let $g : \mathbb{R} \to \mathbb{R}_+$ denote a continuous American payoff satisfying $g(x) \leq C(1 + |x|^k)$ for some constants $C, k \in \mathbb{R}_{++}$. Suppose there exists a European payoff $f : \mathbb{R} \to \mathbb{R}_+$ representing g relative to some $(T_0, x_0) \in C_T$, then:*

1. *The American value function $v_{\mathrm{am},g}$ is analytic on the interior of the set $C_{(T_0,x_0)}$. Moreover, for any $(\widetilde{T}, \widetilde{x}) \in C_{(T_0,x_0)}$ we have $f \overset{(\widetilde{T},\widetilde{x})}{\longrightarrow} g$.*

2. *The representing function is unique up to a Lebesgue nullset, i.e. if* $\tilde{f} \overset{(T_0,x_0)}{\longrightarrow} g$ *then* $\tilde{f} = f$.

3. *The representing function corresponds to the CDEO of* g *relative to* T_0, x_0*, i.e.*

$$f = \mathrm{eu}_{T_0,x_0}(g).$$

4. *For any* $x \in \operatorname{cl} \pi(C_{(T_0,x_0)})$ *we have* $g(x) = \mathrm{am}_{T_0}(f)(x)$ *and hence*

$$g(x) = \mathrm{am}_{T_0}(\mathrm{eu}_{T_0,x_0}(g))(x).$$

5. *The set* $C_{(T_0,x_0)}$ *constitutes a connected component of the continuation region*

$$C'_{T_0} := \{(\vartheta, x) \in [0, T_0] \times \mathbb{R} \mid \mathrm{am}_{T_0}(f)(x) < v_{\mathrm{am},\mathrm{am}_{T_0}(f)}(\vartheta, x)\}$$

associated to the American payoff $\mathrm{am}_{T_0}(f)$*. We have* $f \overset{(T_0,x_0)}{\longrightarrow} \mathrm{am}_{T_0}(f)$ *and therefore*

$$f = \mathrm{eu}_{T_0,x_0}(\mathrm{am}_{T_0}(f)).$$

6. *Suppose that* $\tilde{g} : \mathbb{R} \to \mathbb{R}_+$ *is a continuous American payoff function such that* $\tilde{g} \le g$ *and* $\tilde{g}(x) = g(x)$ *for any* $x \in \operatorname{cl} \pi(C_{(T_0,x_0)})$*. The set* $C_{(T_0,x_0)}$ *is a connected component of the continuation region*

$$\tilde{C}_{T_0} := \{(\vartheta, x) \in [0, T_0] \times \mathbb{R} \mid \tilde{g}(x) < v_{\mathrm{am},\tilde{g}}(\vartheta, x)\}$$

associated to the American payoff $\tilde{g}$ *and*

$$f \overset{(T_0,x_0)}{\longrightarrow} \tilde{g}.$$

Proof.

1. We have $v_{\mathrm{eu},f}(T_0, x_0) = v_{\mathrm{am},g}(T_0, x_0) < \infty$ and therefore Lemma 5.3 implies that the mapping $v_{\mathrm{eu},f}$ is analytic on an open $\mathbb{C}^2$-domain containing the set $(0, T_0) \times \mathbb{R}$. By assumption, the value functions $v_{\mathrm{eu},f}$ and $v_{\mathrm{am},g}$ coincide on $C_{(T_0,x_0)}$. The other assertion is obvious as $C_{(\tilde{T},\tilde{x})}$ is a subset of $C_{(T_0,x_0)}$.

2. Assume that f and $\tilde{f}$ represent g relative to T_0, x_0. Clearly, we have $v_{\mathrm{eu},f}(\vartheta, x) = v_{\mathrm{eu},\tilde{f}}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta, x) < \infty$ for any $(\vartheta, x) \in C_{(T_0,x_0)}$. Lemma 5.3 implies that the value functions $v_{\mathrm{eu},f}$ and $v_{\mathrm{eu},\tilde{f}}$ are analytic on some $\mathbb{C}^2$-domain containing the set $(0, T_0) \times \mathbb{R}$. The set $C_{(T_0,x_0)}$ certainly contains an open ball B. First, we apply the identity theorem to the ϑ variable which shows that the mappings $v_{\mathrm{eu},f}$ and $v_{\mathrm{eu},\tilde{f}}$ coincide on the open stripe $(0, T_0) \times \pi(B)$. Then we apply the identity theorem to the x variable which yields $v_{\mathrm{eu},f}(\vartheta, x) = v_{\mathrm{eu},\tilde{f}}(\vartheta, x) < \infty$ for any $(\vartheta, x) \in (0, T_0) \times \mathbb{R}$. Consequently, it is easy to see that the functions

$$u(y) := \mathrm{N}\left(x_0 + \hat{r}\vartheta_0, \sigma^2\vartheta_0, y\right) f(y),$$

$$\tilde{u}(y) := \mathrm{N}\left(x_0 + \hat{r}\vartheta_0, \sigma^2\vartheta_0, y\right)\tilde{f}(y),$$

where $\vartheta_0 := T_0/2$ and $\hat{r} := r - \sigma^2/2$, are both contained in $\mathrm{L}_1(\mathbb{R})$. Lemma 5.2 yields

$$\begin{aligned} v_{\mathrm{eu},f}(\vartheta_0/2, x/2) &= \int_{\mathbb{R}} \frac{\mathrm{N}\left(x/2 + \hat{r}\vartheta_0/2, \sigma^2\vartheta_0/2, y\right)}{\mathrm{N}\left(x_0 + \hat{r}\vartheta_0, \sigma^2\vartheta_0, y\right)} u(y)\mathrm{d}y \\ &= \sqrt{2}\exp\left(\frac{(x_0 - x/2 + \hat{r}\vartheta_0/2)^2}{\sigma^2\vartheta_0}\right)\int_{\mathbb{R}} \exp\left(-\frac{(y - x + x_0)^2}{2\sigma^2\vartheta_0}\right)u(y)\mathrm{d}y \end{aligned}$$

for any $x \in \mathbb{R}$ and the latter equation remains valid after replacing f and u by $\tilde{f}$ and $\tilde{u}$, respectively. The mappings $v_{\mathrm{eu},f}$ and $v_{\mathrm{eu},\tilde{f}}$ coincide on $(0, T_0) \times \mathbb{R}$ and consequently

$$\int_{\mathbb{R}} \mathrm{N}\left(x_0, \sigma^2\vartheta_0, x - y\right)u(y)\mathrm{d}y = \int_{\mathbb{R}} \mathrm{N}\left(x_0, \sigma^2\vartheta_0, x - y\right)\tilde{u}(y)\mathrm{d}y$$

holds true for any $x \in \mathbb{R}$. We multiply both sides of the latter equation by $e^{\mathrm{i}zx}$, $z \in \mathbb{R}$ and integrate the x variable over the real line. After a few simplifications we obtain

$$\int_{\mathbb{R}} e^{\mathrm{i}zy}u(y)\mathrm{d}y = \int_{\mathbb{R}} e^{\mathrm{i}zy}\tilde{u}(y)\mathrm{d}y$$

for any $z \in \mathbb{R}$. The injectivity of the Fourier transform on $\mathrm{L}_1(\mathbb{R})$ yields that u and $\tilde{u}$ and therefore f and $\tilde{f}$ coincide up to set of zero Lebesgue measure.

3. Clearly, the function f is contained in $\mathrm{eu}_{T_0,x_0}(g)$. It remains to be shown that the latter set is a singleton. For this purpose choose a function $h \in \mathrm{eu}_{T_0,x_0}(g)$ and note that $v_{\mathrm{eu},h}(T_0, x_0) = v_{\mathrm{eu},f}(T_0, x_0) = v_{\mathrm{am},g}(T_0, x_0) < \infty$. By virtue of Lemma 5.3 the mappings $v_{\mathrm{eu},h}$ and $v_{\mathrm{eu},f}$ are analytic on a $\mathbb{C}^2$-domain containing the set $(0, T_0) \times \mathbb{R}$. Due to the second assertion it is sufficient to show $v_{\mathrm{eu},h}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta, x)$ holds true for any $(\vartheta, x) \in C_{(T_0,x_0)}$. Let N denote the set containing all the points from $C_{(T_0,x_0)}$ where the functions $v_{\mathrm{eu},h}$ and $v_{\mathrm{am},g}$ do not coincide. Furthermore, denote by τ_{T_0} the optimal stopping time from (2.7). For any $t \geq 0$ we obtain

$$\mathbb{E}_{x_0}\left[(\hat{v}_{\mathrm{eu},h} - \hat{v}_{\mathrm{am},g})\left(T_0 - t \wedge \tau_{T_0}, X_{t\wedge\tau_{T_0}}\right)\right] = \hat{v}_{\mathrm{eu},h}\left(T_0, x_0\right) - \hat{v}_{\mathrm{am},g}\left(T_0, x_0\right) = 0.$$

The first equality follows from the fact that the discounted European value process as well as the optimally stopped Snell envelope associated to the discounted exercise price process are martingales, cf. [PS, Theorem 2.4 and Remark 2.6]. From Proposition 2.3 we know that $\hat{v}_{\mathrm{eu},h} \geq \hat{v}_{\mathrm{am},g}$ on $[0, T_0] \times \mathbb{R}$ and therefore

$$\mathbb{P}_{x_0}\left((T_0 - t \wedge \tau_{T_0}, X_{t\wedge\tau_{T_0}}) \in N\right) = 0$$

holds true for any $t \geq 0$. Consequently, the set N must have an empty interior which implies that the mappings $v_{\mathrm{eu},h}$ and $v_{\mathrm{am},g}$ coincide on $C_{(T_0,x_0)}$.

4. Choose any $x \in \pi(C_{(T_0,x_0)})$ and pick a $\vartheta_C \in (0, T_0]$ such that $(\vartheta_C, x) \in C_{(T_0,x_0)}$. Due to compactness, we can pick the largest $\vartheta_S \in [0, \vartheta_C)$ such that (ϑ_S, x) is contained in the stopping set. In light of [LB, Theorem 4.1.1], we find that the mapping $v_{\mathrm{am},g}$ is continuous and therefore

$$g(x) \leq \mathrm{am}_{T_0}(f)(x) \leq \liminf_{\vartheta \searrow \vartheta_S} v_{\mathrm{eu},f}(\vartheta, x) = \liminf_{\vartheta \searrow \vartheta_S} v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta_S, x) = g(x).$$

This proves the assertion for $x \in \pi(C_{(T_0,x_0)})$. Next, choose any $x_b \in \partial\pi(C_{(T_0,x_0)})$ and observe that there exists no $\vartheta \in [0, T_0]$ such that (ϑ, x_b) is contained in $C_{(T_0,x_0)}$. Hence, we can find some $\vartheta_b \in (0, T_0]$ such that (ϑ_b, x_b) is located on the boundary of the set $C_{(T_0,x_0)}$. Clearly, there exists an approximating sequence $C_{(T_0,x_0)} \ni (\vartheta_n, x_n) \to (\vartheta_b, x_b)$ as $n \to \infty$ and consequently

$$g(x_b) = \liminf_{n\to\infty} v_{\mathrm{am},g}(\vartheta_n, x_n) = \liminf_{n\to\infty} v_{\mathrm{eu},f}(\vartheta_n, x_n).$$

Applying Fatou's lemma we obtain

$$\begin{aligned}\liminf_{n\to\infty} v_{\mathrm{eu},f}(\vartheta_n, x_n) &\geq v_{\mathrm{eu},f}(\vartheta_b, x_b) \\ &\geq \mathrm{am}_{T_0}(f)(x_b) \geq g(x_b)\end{aligned}$$

and this finally yields $g(x_b) = \mathrm{am}_{T_0}(f)(x_b)$.

5. Clearly, the European payoff f superreplicates $\mathrm{am}_{T_0}(f)$ up to time T_0. Owing to Proposition 2.3, we have $g(x) \leq \mathrm{am}_{T_0}(x)$ and

$$v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{am},\mathrm{am}_{T_0}(f)}(\vartheta, x) \leq v_{\mathrm{eu},f}(x) \tag{2.11}$$

for any $(\vartheta, x) \in [0, T_0] \times \mathbb{R}$. Moreover, equality in (2.11) must hold on the set $C_{(T_0,x_0)}$, as the payoff f represents g relative to T_0, x_0. For any $(\vartheta, x) \in C_{(T_0,x_0)}$ the fourth assertion warrants that $g(x) = \mathrm{am}_{T_0}(f)(x)$ and therefore

$$\mathrm{am}_{T_0}(f)(x) = g(x) < v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{am},\mathrm{am}_{T_0}(f)}(\vartheta, x).$$

This shows that $C_{(T_0,x_0)}$ is a connected subset of C'_{T_0}. Now pick any boundary point $(\vartheta, x) \in \partial C_{(T_0,x_0)}$ with $\vartheta > 0$. Obviously, we have $g(x) = v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x)$. In light of (2.11), we obtain

$$v_{\mathrm{am},\mathrm{am}_{T_0}(f)}(\vartheta, x) \leq v_{\mathrm{eu},f}(x) = g(x) \leq \mathrm{am}_{T_0}(f)(x)$$

which shows that (ϑ, x) is located in the stopping region associated to the American payoff $\mathrm{am}_{T_0}(f)$. In conclusion, we verified that the set $C_{(T_0,x_0)}$ is indeed a connected component of C'_{T_0} and that $\mathrm{am}_{T_0}(f)$ is represented by f relative to T_0, x_0.

6. Choose any $(\vartheta, x) \in C_{(T_0,x_0)}$ and denote by τ_ϑ the optimal stopping time from (2.7). Due to the fact that $X_{\tau_\vartheta} \in \mathrm{cl}\,\pi(C_{(T_0,x_0)})$, we conclude that

$$v_{\mathrm{am},g}(\vartheta, x) = \mathbb{E}_x[e^{-r\tau_\vartheta} g(X_{\tau_\vartheta})] = \mathbb{E}_x[e^{-r\tau_\vartheta} \tilde{g}(X_{\tau_\vartheta})] \leq v_{\mathrm{am},\tilde{g}}(\vartheta, x).$$

The reverse inequality follows immediately from the assumption $\tilde{g} \leq g$ and therefore

$$\tilde{g}(x) = g(x) < v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x) = v_{\mathrm{am},\tilde{g}}(\vartheta, x)$$

holds true. This shows that $C_{(T_0,x_0)}$ is a connected subset of $\tilde{C}_{T_0}$. Next, choose any boundary point $(\vartheta, x) \in \partial C_{(T_0,x_0)}$ and an approximating sequence $(\vartheta_n, x_n)_{n\in\mathbb{N}} \subset C_{(T_0,x_0)}$, i.e. $(\vartheta_n, x_n) \to (\vartheta, x)$ as $n \to \infty$. We have $g(x) = \tilde{g}(x)$ and from above we know that $v_{\mathrm{am},\tilde{g}}(\vartheta_n, x_n) = v_{\mathrm{am}g}(\vartheta_n, x_n)$ for any $n \in \mathbb{N}$. This yields

$$v_{\mathrm{am},\tilde{g}}(\vartheta, x) \leq \liminf_{n\to\infty} v_{\mathrm{am},\tilde{g}}(\vartheta_n, x_n) = \liminf_{n\to\infty} v_{\mathrm{am},g}(\vartheta_n, x_n) = g(x) = \tilde{g}(x)$$

and consequently (ϑ, x) is located within the stopping region associated to the American payoff $\tilde{g}$. Summing up, we have shown that $C_{(T_0,x_0)}$ indeed constitutes a connected component of the set $\tilde{C}_{T_0}$ and that $\tilde{g}$ is represented by f relative to (T_0, x_0).

□

Let us remark that the latter proposition can be extended in many ways. For instance, the reader easily verifies that the second assertion does not explicitly depend on the continuity or integrability properties of the American claim. The key argument relies on the fact that we can choose some open subset of the continuation region where the European value function $v_{\mathrm{eu},f}$ and $v_{\mathrm{eu},\tilde{f}}$ coincide. A more general formulation can be found in Proposition 3.2.

Suppose that we have a finite time horizon $T \in \mathbb{R}_{++}$ and let $f : \mathbb{R} \to \mathbb{R}_+$ denoted some *continuous* European payoff. A key contribution of [JM1] is a sufficient criterion which warrants that the embedded American option $\mathrm{am}_T(f)$ is represented by its generating payoff function f. The upcoming Proposition 2.16 generalizes [JM1, Theorem 3]. First, we fix some notation and prove an auxiliary Lemma.

2.12 Lemma: *Suppose that $f : \mathbb{R} \to \mathbb{R}_+$ is a continuous payoff satisfying $v_{\mathrm{eu},f}(T + \delta, y) < \infty$ for some $y \in \mathbb{R}$ and $\delta > 0$. Moreover, let $g : \mathbb{R} \to \mathbb{R}_+$ denote an upper semi-continuous American payoff which is superreplicated by f up to time T. Then:*

1. *The European value function $v_{\mathrm{eu},f}$ is analytic on an open $\mathbb{C}^2$-domain containing the set $(0, T + \delta) \times \mathbb{C}$ and the embedded American option $\mathrm{am}_T(f)$ is upper semi-continuous.*

2. *The set*
$$M := \{(\vartheta, x) \in [0, T] \times \mathbb{R} \mid g(x) = v_{\mathrm{eu},f}(\vartheta, x)\} \tag{2.13}$$
is closed. For any $x \in \mathbb{R}$ the cut
$$M_x := \{\vartheta \in [0, T] \mid g(x) = v_{\mathrm{eu},f}(\vartheta, x)\} \tag{2.14}$$
is compact. In case that $g = \mathrm{am}_T(f)$, the latter sets are not empty.

3. *For any $\vartheta \in [0,T]$ the first entry time*

$$\tau_\vartheta^M := \inf\{t \in \mathbb{R}_+ \mid (\vartheta - t, X_t) \in M\} \wedge \vartheta \tag{2.15}$$

   is a stopping time.

4. *We have $M = [0,T] \times \mathbb{R}$ if and only if there exist some constants $\gamma_1, \gamma_2 \in \mathbb{R}_+$ such that $f(x) = g(x) = \gamma_1 e^x + \gamma_2 e^{-2r/\sigma^2 x}$ for any $x \in \mathbb{R}$.*

Proof.

1. The desired analyticity follows directly from Lemma 5.3. In particular, for any $\vartheta \in [0,T]$ we find that the mapping $x \mapsto v_{\mathrm{eu},f}(\vartheta, x)$ is a continuous. Proposition 2.3 now implies that $\mathrm{am}_T(f)$ is upper semi-continuous.

2. We note that the mapping $v_{\mathrm{eu},f}$ extends continuously to the set $[0, T+\delta) \times \mathbb{R}$. Obviously, the mapping $v_{\mathrm{eu},f} - g$ is non-negative and lower semi-continuous. This implies that the set $M = \{(\vartheta, x) \in [0,T] \times \mathbb{R} \mid v_{\mathrm{eu},f}(\vartheta, x) - g(x) \leq 0\}$ is closed. The compactness of M_x is evident. In case that $g = \mathrm{am}_T(f)$ holds true, we find that for any $x \in \mathbb{R}$ the infimum in (2.2) is attained at some maturity $\vartheta \in [0,T]$ and therefore M_x and M are not empty.

3. This follows directly from the continuity of the trajectories of X and the closedness of the set M.

4. First, suppose that $f(x) = g(x) = \gamma_1 e^x + \gamma_2 e^{-2r/\sigma^2 x}$ for some constants $\gamma_1, \gamma_2 \in \mathbb{R}_+$. An elementary calculation shows that the European payoff f solves the differential equation $\mathcal{A}f = 0$, where $\mathcal{A} := (r - \frac{\sigma^2}{2})\partial_x + \frac{\sigma^2}{2}\partial_{xx} - r$. Ito's formula implies that $v_{\mathrm{eu},f}(\vartheta, x) = f(x)$ for any $(\vartheta, x) \in \mathbb{R}_+ \times \mathbb{R}$ and therefore $M = [0,T] \times \mathbb{R}$ holds true.

   In order to prove the reverse implication, we assume that $M = [0,T] \times \mathbb{R}$. The latter implies that $g(x) = f(x) = v_{\mathrm{eu},f}(\vartheta, x)$ holds true for any $(\vartheta, x) \in [0,T] \times \mathbb{R}$. Taking the first assertion into account, we conclude that the European payoff f is analytic on an open complex domain containing the real line. Applying Kolmogorov's backward equation yields $\mathcal{A}f = \mathcal{A}v_{\mathrm{eu},f} = \partial_\vartheta v_{\mathrm{eu},f} = 0$. From the theory of ordinary differential equations we know that f can be represented as a linear combination of the two fundamental solutions e^x and $e^{-2r/\sigma^2 x}$. The payoff f is assumed to be non-negative and consequently there exist constants $\gamma_1, \gamma_2 \in \mathbb{R}_+$ such that $f(x) = \gamma_1 e^x + \gamma_2 e^{-2r/\sigma^2 x}$ for any $x \in \mathbb{R}$.

$\square$

Let us remark that the latter lemma can be readily generalized to include market models which are driven by other continuous diffusion processes than the geometric Brownian motion. The first assertion requires certain analyticity properties of the Markov transition kernel, see the proof of Lemma 5.3. The fourth assertion can be restated in terms of the invariant functions associated to the pricing semi-group. For the subsequent proposition, please recall our convention $\sup \emptyset := -\infty$.

2.16 Proposition: *Assume that $f : \mathbb{R} \to \mathbb{R}_+$ is a continuous European payoff satisfying $v_{\text{eu},f}(T+\delta, y) < \infty$ for some $y \in \mathbb{R}$ and $\delta > 0$. Moreover, let $g : \mathbb{R} \to \mathbb{R}_+$ denote an upper semi-continuous American payoff which is superreplicated by f up to time T. Suppose that the quantities M, M_x, τ_ϑ^M are defined as in Lemma 2.12 and that C_T denotes the continuation set associated to g as defined in* (1.7)*. Then:*

1. *For any $(\vartheta, x) \in [0, T] \times \mathbb{R}$ such that*
$$H_M(\vartheta, x) := \mathbb{P}_x\left((\vartheta - \tau_\vartheta^M, X_{\tau_\vartheta^M}) \in M\right) = 1 \tag{2.17}$$
we have $v_{\text{am},g}(\vartheta, x) = v_{\text{eu},f}(\vartheta, x)$.

2. *We have*
$$\begin{gathered}\{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid \sup M_x < \vartheta \text{ and } H_M(\vartheta, x) = 1\} \\ \cap \\ C_T \\ \cap \\ \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid \sup M_x < \vartheta\}.\end{gathered}$$
In particular, the set M is contained in the stopping region associated to g.

Proof.

1. The European payoff f superreplicates g up to T and owing to Proposition 2.3 the inequality $v_{\text{am},g} \leq v_{\text{eu},f}$ holds true on $[0, T] \times \mathbb{R}$. Now choose any $(\vartheta, x) \in [0, T] \times \mathbb{R}$ such that $H_M(\vartheta, x) = 1$. The log-price process X has continuous trajectories and therefore $\mathbb{P}_x\left(g(X_{\tau_\vartheta^M}) = v_{\text{eu},f}(\vartheta - \tau_\vartheta^M, X_{\tau_\vartheta^M})\right) = 1$. In Lemma 2.12 it was shown that τ_ϑ^M is a finite stopping time. The discounted European value process is a martingale and consequently we obtain
$$\begin{aligned} v_{\text{am},g}(\vartheta, x) &\geq \mathbb{E}_x\left[e^{-r\tau_\vartheta^M} g(X_{\tau_\vartheta^M})\right] \\ &= \mathbb{E}_x\left[e^{-r\tau_\vartheta^M} v_{\text{eu},f}(\vartheta - \tau_\vartheta^M, X_{\tau_\vartheta^M})\right] \\ &= v_{\text{eu},f}(\vartheta, x) \end{aligned} \tag{2.18}$$
by optional sampling. This proves the assertion.

2. In order to verify the first inclusion, choose any $(\vartheta, x) \in (0, T] \times \mathbb{R}$ such that $\sup M_x < \vartheta$ and $H_M(\vartheta, x) = 1$. The definition of the set M_x implies that $g(x) < v_{\text{eu},f}(\vartheta, x)$. Moreover, by virtue of the first assertion we have $v_{\text{eu},f}(\vartheta, x) = v_{\text{am},g}(\vartheta, x)$ which shows that (ϑ, x) is indeed contained in the continuation region C_T.

   For the proof of the second inclusion, choose any $(\vartheta, x) \in C_T$, i.e. we have $g(x) < v_{\text{am},g}(\vartheta, x) \leq v_{\text{eu},f}(\vartheta, x)$. In case that the set M_x is empty, we can directly conclude that $\sup M_x = -\infty < \vartheta$ holds true. If the compact set M_x is not empty, we have $\sup M_x \in [0, T]$. Taking the monotonicity properties of the American value function into account, we obtain $g(x) = v_{\mathrm{am},g}(\vartheta', x) = v_{\mathrm{eu},f}(\sup M_x, x)$ for any $\vartheta' \in [0, \sup M_x]$. This shows that $\vartheta > \sup M_x$ must indeed hold true which proves the assertion.

□

The quantity $H_M(\vartheta, x)$ defined in (2.17) corresponds to the probability that the space-time process started at time $T - \vartheta$ and spot x hits the zero set of the mapping $v_{\mathrm{eu},f} - g$. We notice that the validity of Proposition 2.16 does not specifically depend on the distributional properties of the geometric Brownian motion. The reader may adapt the results if the stock price dynamics are modeled by a different, possibly multivariate, continuous diffusion process.
Let us add some further remarks: Owing to the first assertion of Lemma 2.12, we conclude that Proposition 2.16 is in particular applicable to the embedded American option $\mathrm{am}_T(f)$ associated to the European payoff f. Example 2.25 will show that the second assertion of the latter proposition is meaningful, i.e. the three sets do not trivially coincide. Furthermore, if there exists a continuous mapping $\widehat{\vartheta} : \mathbb{R} \to [0, T]$ such that the graph of $\widehat{\vartheta}$ is a subset of M, we can obtain [JM1, Theorem 5] as a special case of Proposition 2.16, cf. Assertion 2 of Proposition 2.19. In Example 2.31 we will see that the continuity assumption of the Jourdain-Martini theorem is indeed a bit too restrictive and therefore Proposition 2.16 earns the right to exist. Besides, the latter example sheds light on the fact that American claims with spatially discontinuous early exercise boundaries need to be considered when searching for a general characterization of all representable American options. We call a function $b : \mathbb{R} \to [0, T]$ *early exercise boundary* or *stopping boundary* if $C_T = \{(\vartheta, x) \in [0, T] \times \mathbb{R} \mid \vartheta > b(x)\}$. The curve b can possess certain continuity and smoothness properties, e.g. for the American put it can be shown that the associated stopping boundary is continuous and increasing, cf. [PS, Theorem 25.3]. In case that b is a discontinuous function we say that the early exercise boundary is *spatially discontinuous*. Let us remark that there is no standardized definition of the early exercise boundary in the literature. For instance, some authors prefer a parametrization with respect to time, cf. [PS, Equation 25.2.8].

2.19 Proposition: *Assume that f is a continuous European payoff such that $v_{\mathrm{eu},f}(T + \delta, y) < \infty$ for some $y \in \mathbb{R}$ and $\delta > 0$. Furthermore, let $g : \mathbb{R} \to \mathbb{R}_+$ denote an American payoff which is superreplicated by f up to time T. Suppose there exists a continuous mapping $\widehat{\vartheta} : \mathbb{R} \to [0, T]$ such that*

$$g(x) = v_{\mathrm{eu},f}(\widehat{\vartheta}(x), x)$$

for any $x \in \mathbb{R}$. Moreover, we write

$$\widetilde{C}_T := \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid \vartheta > \widehat{\vartheta}(x)\}.$$

Let π and M_x be defined as in (2.9) and (2.14), respectively. Then:

1. *The American option g is continuous.*

2. *For any $(\vartheta, x) \in \widetilde{C}_T$ we have $v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x)$, cf. [JM1, Theorem 5].*

3. *The continuation set can be represented as follows:*

$$C_T = \widetilde{C}_T \cap \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid \max M_x < T\} \tag{2.20}$$

Moreover, for any $x \in \pi(C_T)$ we have $\widehat{\vartheta}(x) = \max M_x$.

4. *For any $\vartheta \in [0, T]$ the stopping time*

$$\tilde{\tau}_\vartheta := \inf \{t \in \mathbb{R}_+ \mid \vartheta - t \leq \widehat{\vartheta}(X_t)\} \wedge \vartheta \tag{2.21}$$

is optimal for the stopping problem inherent in (1.6).

5. *The mapping $\mathbb{R} \ni x \mapsto \max M_x$ parametrizes the early exercise boundary associated to the American payoff g, i.e. $C_T = \{(\vartheta, x) \in [0, T] \times \mathbb{R} \mid \vartheta > \max M_x\}$. The curve $\widehat{\vartheta}$ differs from the early exercise boundary only at spot prices where immediate exercising the American option is optimal, i.e. $\widehat{\vartheta}(x) < \max M_x$ implies $M_x = [0, T]$. Consequently, the European payoff f globally represents g up to time T.*

6. *On the set*

$$G := \left\{x \in \mathbb{R} \mid 0 < \widehat{\vartheta}(x) < T \text{ and } g \text{ is } C^2 \text{ on some neighborhood containing } x\right\}$$

the American payoff g satisfies

$$(\mathcal{A}g)(x) \leq 0. \tag{2.22}$$

Here we denote by $\mathcal{A} := (r - \frac{\sigma^2}{2})\partial_x + \frac{\sigma^2}{2}\partial_{xx} - r$ the infinitesimal generator associated to the pricing semi-group $Q_\vartheta[h](x) := \mathbb{E}_x[e^{-r\vartheta} h(X_\vartheta)]$ on C^2.

Proof.

1. Lemma 2.12 warrants that the mapping $v_{\mathrm{eu},f}$ is continuous on $[0, T + \delta) \times \mathbb{R}$. The curve $\widehat{\vartheta}$ was assumed to be continuous and therefore the payoff g is continuous as well.

2. Suppose that M and H_M are defined as (2.13) and (2.17), respectively. Obviously, the graph of the curve $\widehat{\vartheta}$ is a subset of M. The mapping $\widehat{\vartheta}$ was assumed to be continuous and this shows that $H_M(\vartheta, x) = 1$ for any $(\vartheta, x) \in (0, T] \times \mathbb{R}$ such that $\widehat{\vartheta}(x) < \vartheta$. Proposition 2.16 now directly implies the assertion.

3. Suppose that $\widehat{\vartheta}(x) < \max M_x$ for some $x \in \mathbb{R}$. The second assertion of the proposition yields $v_{\mathrm{eu},f}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta, x)$ for any maturity $\vartheta \in (\widehat{\vartheta}(x), \max M_x]$. The American value function is increasing in the first variable and consequently

$$v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{am},g}(\max M_x, x) = v_{\mathrm{eu},f}(\max M_x, x) = g(x).$$

This shows that $v_{\mathrm{eu},f}(\vartheta, x) = g(x)$ for any $\vartheta \in [\widehat{\vartheta}(x), \max M_x]$. By virtue of Lemma 2.12 and the identity theorem we obtain $v_{\mathrm{eu},f}(\vartheta, x) = g(x)$ for any $\vartheta \in$

$[0,T]$ and hence $\max M_x = T$. This finally shows that $v_{\mathrm{am},g}(\vartheta, x) = g(x)$ for any $\vartheta \in [\widehat{\vartheta}(x), T]$ and therefore $x \notin \pi(C_T)$.

In order to verify the first inclusion of Equation (2.20), choose any $(\vartheta, x) \in C_T$. We note that $\max M_x < T$ holds true. Indeed, assuming the opposite yields the contradiction $v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{am},g}(T, x) \leq v_{\mathrm{eu},f}(T, x) = g(x)$. In the same manner we can conclude that $\vartheta > \widehat{\vartheta}(x)$. Indeed, assuming that $\vartheta \leq \widehat{\vartheta}(x)$ implies $g(x) < v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{am},g}(\widehat{\vartheta}(x), x) \leq v_{\mathrm{eu},f}(\widehat{\vartheta}(x), x) = g(x)$ and this is clearly not possible. In conclusion we have shown that $C_T \subset \widetilde{C}_T \cap \{(\vartheta, x) \in (0,T] \times \mathbb{R} \mid \max M_x < T\}$.

To verify the reverse inclusion, choose any $(\vartheta, x) \in \widetilde{C}_T$ such that $\max M_x < T$. Due to the monotonicity of the American value function in the first variable, we find that $\widehat{\vartheta}(x) = \max M_x$. Indeed, assuming the opposite yields $v_{\mathrm{am},g}(\vartheta', x) \leq v_{\mathrm{am},g}(\max M_x, x) \leq v_{\mathrm{eu},f}(\max M_x, x) = g(x)$ for any $\vartheta' \in (\widehat{\vartheta}(x), \max M_x)$. Owing to Lemma 2.12 and the identity theorem, we then conclude that $v_{\mathrm{eu},f}(\vartheta', x) = g(x)$ holds true for any $\vartheta' \in [0,T]$ and therefore $\max M_x = T$. This clearly contradicts our assumption. Consequently, we find that $\vartheta > \widehat{\vartheta}(x) = \max M_x$ holds true. In light of the second assertion and the definition of the set M_x, we obtain $v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},f}(\vartheta, x) > g(x)$ and this shows that (ϑ, x) is indeed contained in C_T.

4. Literally the same calculation as in (2.18) yields the optimality of the stopping time from (2.21).

5. Suppose that $\widehat{\vartheta}(x) < \max M_x$ for some $x \in \mathbb{R}$. Owing to Assertion 3, we have $x \notin \pi(C_T)$ and a simple calculation yields $M_x = [0,T]$, for details see Lemma 2.24 below. From Proposition 2.16 we know that M_x is contained in the stopping region associated to g. We conclude that the mappings $\widehat{\vartheta}$ and $x \mapsto \max M_x$ differ only at spot prices where immediate exercising is optimal. Moreover, Equation (2.20) shows that $\mathbb{R} \ni x \mapsto \max M_x$ indeed parametrizes the early exercise boundary associated to g. From (2.20) it is obvious that $C_T \subset \widetilde{C}_T$ and taking the second assertion into account, we find that the European payoff f globally represents g up to time T in the sense of Definition 2.8.

6. Clearly, the mapping $\Psi := v_{\mathrm{eu},f} - g$ is C^2 on the set $(0,T) \times G$ and there we have

$$
\begin{aligned}
\mathcal{A}g &= \mathcal{A}v_{\mathrm{eu},f} - \mathcal{A}\Psi \\
&= \partial_\vartheta v_{\mathrm{eu},f} - \left(r - \frac{\sigma^2}{2}\right)\partial_x \Psi - \frac{\sigma^2}{2}\partial_{xx}\Psi - r\Psi \\
&= c\,\nabla\Psi - \frac{\sigma^2}{2}\partial_{xx}\Psi - r\Psi
\end{aligned}
\tag{2.23}
$$

where $c := (1, \frac{\sigma^2}{2} - r)$. Now choose any $x \in G$. By definition we have $\Psi(\widehat{\vartheta}(x), x) = 0$. Due to the fact that Ψ only assumes non-negative values, the first order condition $(\nabla\Psi)(\widehat{\vartheta}(x), x) = 0$ and the second order condition $(\partial_{xx}\Psi)(\widehat{\vartheta}(x), x) \geq 0$ hold true. From (2.23) we obtain

$$(\mathcal{A}g)(x) = -\frac{\sigma^2}{2}(\partial_{xx}\Psi)(\widehat{\vartheta}(x), x) \leq 0$$

which concludes the proof.

□

Let us add a few comments concerning Proposition 2.19: Besides the stopping time $\tilde{\tau}_\vartheta$ from (2.21) there may exist other stopping times which are optimal for the stopping problem inherent in (1.6). For instance, the first entry time τ_ϑ into the stopping set, as defined in (2.7), is an optimal choice, cf. [PS, Theorem 2.7]. Equation (2.20) shows that $\tilde{\tau}_\vartheta$ and τ_ϑ do not necessarily coincide. Moreover, if the sets from (2.14) satisfy $|M_x| = 1$ for any $x \in \mathbb{R}$, it is not hard to show that the mapping $\mathbb{R} \ni x \mapsto \max M_x$ is continuous, cf. [JM1, Remark 4]. Here it is crucial that the European payoff f is assumed to be continuous. Indeed, in Example 2.25 we will consider a discontinuous European payoff which generates a discontinuous curve of unique minimal points and a discontinuous EAO g which is not represented by its generating European claim. Interestingly, the American option from the latter example is locally representable on any connected component of the associated continuation set but not globally representable by some European payoff.

The following lemma is a little side note to Proposition 2.19. It describes the set of logarithmic spot prices where the mappings $\widehat{\vartheta}$ and $x \mapsto \max M_x$ differ.

2.24 Lemma: *Suppose that we are in the setting of Proposition 2.19 and define*

$$D := \{x \in \mathbb{R} \mid \widehat{\vartheta}(x) < \max M_x\}.$$

Then:

1. *For any $x \in D$ we have $M_x = [0, T]$. To put differently, for any $x \in \mathbb{R}$ such that the mapping $[0, T] \ni \vartheta \mapsto v_{\mathrm{eu},f}(\vartheta, x)$ is not constant, we have $\widehat{\vartheta}(x) = \max M_x$.*

2. *If there exists a continuous mapping $\widehat{\rho} : \mathbb{R} \to [0, T]$ such that $\widehat{\vartheta} \neq \widehat{\rho}$ and $g(x) = v_{\mathrm{eu},f}(\widehat{\rho}(x), x)$ for any $x \in \mathbb{R}$, then the set D clusters at some point.*

3. *If the set D clusters at some point $x_0 \in \mathbb{R}$, then there exist constants $\gamma_1, \gamma_2 \in \mathbb{R}_+$ such that $f(x) = g(x) = \gamma_1 e^x + \gamma_2 e^{-2r/\sigma^2 x}$ for any $x \in \mathbb{R}$ and we have $M = [0, T] \times \mathbb{R}$.*

Proof.

1. Suppose that $x \in D$. The second assertion of Proposition 2.19 yields $v_{\mathrm{eu},f}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta, x)$ for any maturity $\vartheta \in (\widehat{\vartheta}(x), \max M_x]$. The American value function is increasing in the first variable and consequently

$$v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{am},g}(\max M_x, x) = v_{\mathrm{eu},f}(\max M_x, x) = g(x).$$

This shows that $v_{\mathrm{eu},f}(\vartheta, x) = g(x)$ for any $\vartheta \in (\widehat{\vartheta}(x), \max M_x]$. From Lemma 2.12 we know that the mapping $v_{\mathrm{eu},f}$ is analytic on some $\mathbb{C}^2$-domain containing $(0, T+\delta) \times \mathbb{R}$. By virtue of the identity theorem we obtain $v_{\mathrm{eu},f}(\vartheta, x) = g(x)$ for any $\vartheta \in [0, T+\delta)$ and this proves $M_x = [0, T]$.

2. Clearly, we can choose an open interval $I \subset \mathbb{R}$ such that $\widehat{\vartheta}(x) < \widehat{\rho}(x)$ for any $x \in I$ or $\widehat{\vartheta}(x) > \widehat{\rho}(x)$ for any $x \in I$. We only consider the first case as the second follows by interchanging the notation. We have $\widehat{\vartheta}(x) < \widehat{\rho}(x) \leq \max M_x$ for any $x \in I$ and therefore $I \subset D$.

3. Assume that the set D clusters at some point x_0. Choose a sequence $(x_n)_{n\in\mathbb{N}}$ contained in D such that $x_n \to x_0$ as $n \to \infty$. From the first assertion we know that $v_{\mathrm{eu},f}(\vartheta, x_n) = g(x_n)$ and therefore $(\partial_\vartheta v_{\mathrm{eu},f})(\vartheta, x_n) = 0$ holds true for any $\vartheta \in (0, T)$ and $n \in \mathbb{N}$. The identity theorem implies that $\partial_\vartheta v_{\mathrm{eu},f} = 0$ on the set $(0, T) \times \mathbb{R}$. Hence, we obtain $g(x) = f(x) = v_{\mathrm{eu},f}(T/2, x)$ for any $x \in \mathbb{R}$. In other words, the European payoff f is an entire function satisfying the ordinary differential equation $\mathcal{A}f = 0$ where $\mathcal{A}$ is defined as in Assertion 6 of Proposition 2.19. Consequently, there exist constants $\gamma_1, \gamma_2 \in \mathbb{R}_+$ such that $f(x) = \gamma_1 e^x + \gamma_2 e^{-2r/\sigma^2 x}$ for any $x \in \mathbb{R}$. Lemma 2.12 now directly implies that $M = [0, T] \times \mathbb{R}$. □

We want to add some concluding remarks. Many of the results contained in this section can be generalized for larger classes of European payoff functions. Our main results involve the embedding of American claims into measure type European payoffs, cf. Theorem 2.52. Naturally, some ideas for possible extensions in this direction are implicitly contained in the just mentioned theorem and in Section 2.4. For instance, a generalization of the second and third assertion from Proposition 2.10 can be found in Lemma 3.2. Due to the fact that many of the arguments above hardly rely on the Black-Scholes model under consideration, the interested reader may also attempt to transfer the obtained results to other market models. A careful analysis shows that many proofs can easily be adapted to Markovian models with continuous sample paths where any two states of the stock price process communicate.

2.1.2 Examples

This subsection serves different purposes: On the one hand, we want to emphasize that representability is a property which is not only satisfied by some "exotic" American payoffs. In particular, we will verify that the American butterfly in the Bachelier model is representable. On the other hand, we will present some limit cases which hopefully provide the reader with some further insight into the mathematical notions from the preceding sections. For instance, Example 2.25 shows that there exist American options which are locally but not globally representable in the sense of Definition 2.8. Besides, we will see that an embedded American option which is not represented by its generating European payoff still might be representable. Moreover, Example 2.31 indicates that American options with discontinuous early exercise boundaries still might be globally representable. Consequently, one might say that the continuity assumption imposed by

[JM1, Theorem 5] on the curve of minimal points $\widehat{\vartheta}$ is a bit too restrictive. In particular, we may infer that the generalization formulated in Proposition 2.16 is indeed meaningful. These thoughts should be kept in mind when attempting the venture of characterizing all representable American claims.

We reuse the notational conventions from Subsection 2.1.1. In particular, given a European payoff $f : \mathbb{R} \to \mathbb{R}_+$, the associated value function $v_{\mathrm{eu},f}$ is defined as in (1.4). The associated embedded American option up to some time $T \in [0, \infty]$ is denoted by $\mathrm{am}_T(f)$, see (2.2). Furthermore, we write $\varphi := \mathrm{N}(0, 1, \cdot)$ for the probability density function and Φ for the cumulative distribution function of a standard normal random variable.

2.25 Example: Consider the Black-Scholes market (1.3) with $B_0 = 1, r = 1, \sigma = \sqrt{2}$, i.e.

$$\begin{aligned} \mathrm{d}B_t &= B_t \mathrm{d}t, \\ \mathrm{d}X_t &= \sqrt{2}\mathrm{d}W_t, \end{aligned}$$

and fix some time horizon $T \in \mathbb{R}_{++}$. The European value function associated to the payoff $f := \mathbb{1}_{[0,1]}$ is given by

$$v_{\mathrm{eu},f}(\vartheta, x) = e^{-\vartheta} \left(\Phi\left(\frac{1-x}{\sqrt{2\vartheta}}\right) - \Phi\left(-\frac{x}{\sqrt{2\vartheta}}\right)\right).$$

An elementary calculation yields

$$\lim_{\vartheta \searrow 0} v_{\mathrm{eu},f}(\vartheta, x) = \begin{cases} \frac{1}{2} & \text{if } x \in \{0, 1\}, \\ f(x) & \text{otherwise.} \end{cases} \tag{2.26}$$

Moreover, for any $(\vartheta, x) \in \mathbb{R}_{++} \times \mathbb{R}$ we have

$$\partial_\vartheta v_{\mathrm{eu},f}(\vartheta, x) = -e^{-\vartheta}(2\vartheta)^{-\frac{3}{2}} \left(\varphi\left(\frac{1-x}{\sqrt{2\vartheta}}\right)(1-x) + \varphi\left(\frac{x}{\sqrt{2\vartheta}}\right) x \right) - v_{\mathrm{eu},f}(\vartheta, x).$$

The latter equation implies that $\partial_\vartheta v_{\mathrm{eu},f}(\vartheta, x) < 0$ for any $(\vartheta, x) \in \mathbb{R}_{++} \times [0, 1]$. Taking (2.26) into account, we conclude that the embedded American option is given by

$$g(x) := \mathrm{am}_T(f)(x) = v_{\mathrm{eu},f}(T, x)\, \mathbb{1}_{[0,1]}(x).$$

The function g attains its global maximum at $x^* := \frac{1}{2}$ and satisfies $g(x^* + x) = g(x^* - x)$ as well as $g(x) < 2g(0)$ for any $x \in \mathbb{R}$. Figure 2.2 depicts the graph of the embedded American option for different time horizons. Clearly, for any $x \in \mathbb{R}$ the infimum in (2.2) is attained at the unique point

$$\widehat{\vartheta}(x) = T\, \mathbb{1}_{[0,1]}(x).$$

This shows that neither the embedded American option nor the associated curve $\widehat{\vartheta}$ of unique minima need to be continuous if the underlying European payoff was discontinuous at first. The reader may compare this result to the statements of Proposition 2.19

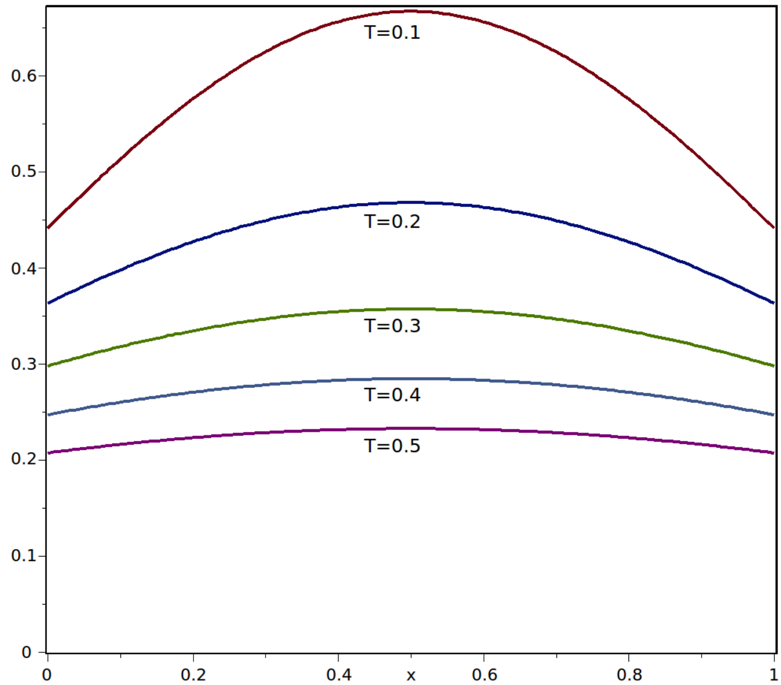

Figure 2.2: The EAO from Example 2.25 for different terminal times T.

and [JM1, Remark 4]. Let C_T denote the continuation set associated to g as defined in (1.7). Proposition 2.16 yields that

$$C_T \subset \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid x \notin [0, 1]\} \,. \tag{2.27}$$

Clearly, for any $(\vartheta, x) \in (0, T] \times \mathbb{R}$ such that $x \notin [0, 1]$ we have $H_M(\vartheta, x) < 1$ where H_M is defined as in (2.17). Hence, Proposition 2.16 does not provide any further information concerning the set C_T or the representability of the American payoff g. For any $\vartheta \in (0, T]$ and $x \notin [0, 1]$ we have $g(x) = 0 < v_{\mathrm{am},g}(\vartheta, x)$ and therefore $(\vartheta, x) \in C_T$. This shows that the reverse inclusion in (2.27) holds true, i.e.

$$C_T = \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid x \notin [0, 1]\} \,.$$

Next, we will show the embedded American payoff g is *not* represented by its generating European claim f. To this end choose any $(\vartheta, x) \in C_T$ and denote by τ_ϑ the stopping time from (2.7) which is optimal for the stopping problem inherent in (1.6), cf. [PS, Theorem 2.7]. From above we know that $\partial_\vartheta v_{\mathrm{eu},f} < 0$ holds on the set $\mathbb{R}_{++} \times [0, 1]$ and consequently we obtain

$$\begin{aligned} v_{\mathrm{am},g}(\vartheta, x) &= \mathbb{E}_x \left[g(X_{\tau_\vartheta}) e^{-r\tau_\vartheta} \right] \\ &= \mathbb{E}_x \left[\mathbb{1}_{[0,1]}(X_{\tau_\vartheta}) v_{\mathrm{eu},f}(T, X_{\tau_\vartheta}) e^{-r\tau_\vartheta} \right] \\ &< \mathbb{E}_x \left[v_{\mathrm{eu},f}(\vartheta - \tau_\vartheta, X_{\tau_\vartheta}) e^{-r\tau_\vartheta} \right] \\ &= v_{\mathrm{eu},f}(\vartheta, x). \end{aligned}$$

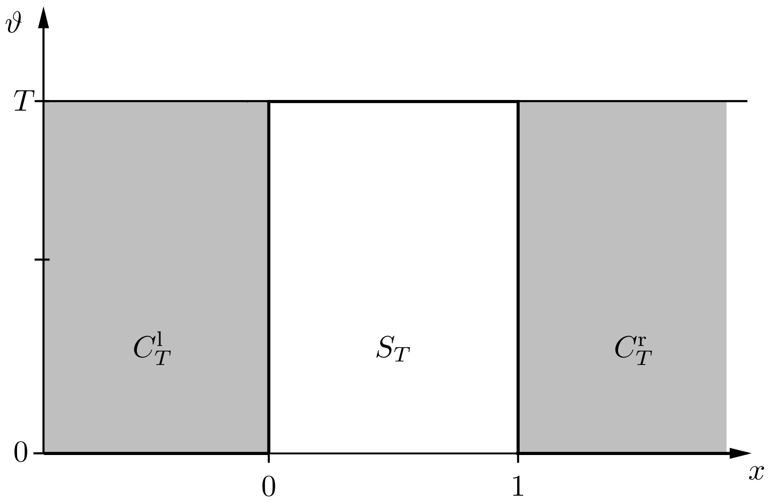

Figure 2.3: The sets $C_T = C_T^{\mathrm{l}} \cup C_T^{\mathrm{r}}$ and S_T associated to g.

The last equality follows from the optional sampling theorem applied to the discounted European value process. We conclude that $v_{\mathrm{am},g}(\vartheta, x) < v_{\mathrm{eu},f}(\vartheta, x)$ for any $(\vartheta, x) \in C_T$ and therefore the payoff g is indeed not represented by f.
Nonetheless, there exist unique European payoff functions which *locally* represent g on the connected components $C_T^{\mathrm{l}} := [T, 0) \times (-\infty, 0)$ and $C_T^{\mathrm{r}} := [T, 0) \times (1, \infty)$ of the continuation set C_T, see Figure 2.3. First, we will verify that $h(x) := 2g(0)\cosh(x)\,\mathbb{1}_{\mathbb{R}_+}(x)$ represents g on the left connected component C_T^{l}. The European value function associated to h is given by

$$v_{\mathrm{eu},h}(\vartheta, x) = 2g(0)e^{-\vartheta} H(\vartheta, x)$$

where $H(\vartheta, x) := \mathbb{E}_x[\cosh(X_\vartheta)\mathbb{1}_{\mathbb{R}_+}(X_\vartheta)]$. For any $x < 0$ we have $v_{\mathrm{eu},h}(0, x) = 0$ and owing to the symmetry of the hyperbolic cosine function, we obtain

$$v_{\mathrm{eu},h}(\vartheta, 0) = 2g(0)e^{-\vartheta}\mathbb{E}_0[\cosh(X_\vartheta)\mathbb{1}_{\mathbb{R}_+}(X_\vartheta)] = g(0)e^{-\vartheta}\mathbb{E}[\cosh(\mathcal{N}_{0,2\vartheta})] = g(0)$$

for any $\vartheta \in (0, T]$. Moreover, applying partial integration twice yields

$$\begin{aligned} H(\vartheta, x) &= \int_0^\infty \cosh(y)\mathrm{N}(x, 2\vartheta, y)\mathrm{d}y \\ &= \int_0^\infty \cosh(y)\partial_{xx}\mathrm{N}(x, 2\vartheta, y)\mathrm{d}y - \partial_x \mathrm{N}(0, 2\vartheta, x) \\ &= \partial_{xx} H(\vartheta, x) + \frac{x}{2\vartheta}\mathrm{N}(0, 2\vartheta, x). \end{aligned}$$

Taking the boundary conditions $H(\vartheta, 0) = \frac{1}{2}e^\vartheta$ and $H(0, x) = 0$ for $x < 0$ into account, we apply Lemma 5.5 in order to derive the explicit representation

$$H(\vartheta, x) = \frac{e^{-x}}{4}\left(e^{2x} - 1 + e^\vartheta\left(\operatorname{erf}\left(\frac{x - 2\vartheta}{2\sqrt{\vartheta}}\right) + e^{2x}\operatorname{erf}\left(\frac{x + 2\vartheta}{2\sqrt{\vartheta}}\right) + 2\right)\right)$$

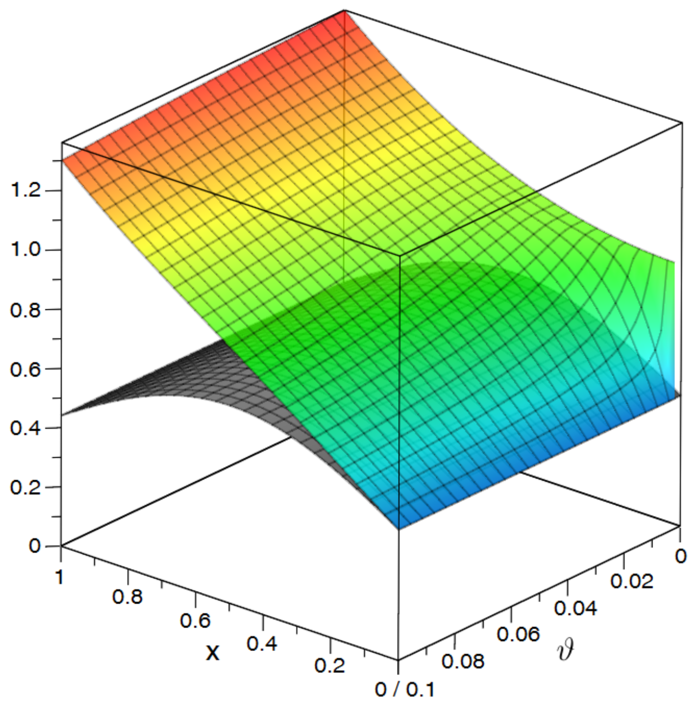

Figure 2.4: The European value function $v_{\mathrm{eu},h}$ and the payoff g for $T = \frac{1}{10}$.

where $\mathrm{erf}(z) := 2\Phi(\sqrt{2}z) - 1$ denotes the Gauss error function. Figure 2.4 depicts the graph of $v_{\mathrm{eu},h}$ for $T = \frac{1}{10}$. Please, note that the x-axis and ϑ-axis are reversed in order to improve the visualization. Let us verify that h superreplicates the American payoff g up to time T. Clearly, we have $h(x) \geq 2g(0)\,\mathbb{1}_{[0,1]}(x) \geq g(x)$ for any $x \in \mathbb{R}$. Thus, it is sufficient to show that $v_{\mathrm{eu},h} \geq g$ on the set $[0,T] \times (0,1]$. An easy calculation yields

$$\partial_\vartheta v_{\mathrm{eu},h}(\vartheta, x) = -\frac{g(0)}{2\sqrt{\pi}\vartheta^{\frac{3}{2}}} e^{-\vartheta - x} \left(e^{\frac{x(4\vartheta - x)}{4\vartheta}} x + \vartheta^{\frac{3}{2}} \sqrt{\pi}(e^{2x} - 1) \right) < 0$$

for any $(\vartheta, x) \in [0,T] \times (0, \infty)$. Hence, we only need to show that the quantity $d(x) := v_{\mathrm{eu},h}(T,x) - g(x)$ is non-negative for any $x \in (0,1]$. The mapping $h - f$ is increasing on $(0,\infty)$ and vanishes on $(-\infty, 0]$. Due to the distributional properties of the Gaussian law, we find that the inequality

$$d(x) = v_{\mathrm{eu},h-f}(T,x) = e^{-T}\mathbb{E}[(h-f)(\mathcal{N}_{x,2T})] \geq e^{-T}\mathbb{E}[(h-f)(\mathcal{N}_{0,2T})] = d(0) = 0$$

holds true for any $x \in (0,1]$. Consequently, the European payoff h indeed superreplicates g up to time T and owing to Proposition 2.3 we have $v_{\mathrm{am},g} \leq v_{\mathrm{eu},h}$ on $[0,T] \times \mathbb{R}$. It was already shown that the functions $v_{\mathrm{eu},h}$ and g coincide on the stopping boundary associated to C_T^{l}, i.e. $v_{\mathrm{eu},h}(0,x) = g(x) = 0$ for any $x < 0$ and $v_{\mathrm{eu},h}(\vartheta, 0) = g(0)$ for any $\vartheta \in [0,T]$. Consequently, for any $(\vartheta, x) \in C_l$ we obtain

$$v_{\mathrm{am},g}(\vartheta, x) = \mathbb{E}_x \left[g(X_{\tau_\vartheta}) e^{-r\tau_\vartheta} \right]$$

$$= \mathbb{E}_x \left[v_{\mathrm{eu},h}(\vartheta - \tau_\vartheta, X_{\tau_\vartheta}) e^{-r\tau_\vartheta} \right]$$
$$= v_{\mathrm{eu},h}(\vartheta, x)$$

by optional sampling, where τ_ϑ denotes the optimal stopping time from (2.7). This finally shows that the American payoff g is represented by h on the left connected component C_T^{l}. Taking the symmetry of the problem into account, the reader readily verifies that $\tilde{h}(x) := 2g(1)\mathbb{1}_{(-\infty,1]}(x)\cosh(x-1)$ represents the embedded American option g on the right connected component C_T^{r} of the continuation set. Proposition 3.2 shows that h and $\tilde{h}$ are uniquely determined.[1]

Summing up, for any time horizon $T \in \mathbb{R}_{++}$ the embedded American option $\mathrm{am}_T(f)$ as well as the associated curve of unique minima are discontinuous and $\mathrm{am}_T(f)$ is not represented by its generating payoff f. Nevertheless, on each connected component of the associated continuation region the American payoff $\mathrm{am}_T(f)$ is locally represented by another, uniquely determined European claim. In particular, this example demonstrates that American payoffs which are representable with respect to any point of the corresponding continuation set do not need to be globally representable in the sense of Definition 2.8. Moreover, we observe that the EAO is piecewise analytic on $\mathbb{R} \setminus \{0, 1\}$.

2.28 Example: We consider the Black-Scholes market

$$\mathrm{d}B_t = rB_t\mathrm{d}t,$$
$$\mathrm{d}X_t = \left(r - \frac{\sigma^2}{2}\right)\mathrm{d}t + \sigma\mathrm{d}W_t,$$

from (1.3) with $B_0 = 1, r \geq 0, \sigma > 0$ and some time horizon $T \in (0, \infty]$. Choose $b_1 \in (-\frac{2r}{\sigma^2}, 1), b_2 \notin (-\frac{2r}{\sigma^2}, 1)$ and define $p(z) := \frac{\sigma^2}{2}z^2 + \left(r - \frac{\sigma^2}{2}\right)z - r$. The value function associated to the European payoff

$$f(x) := p(b_2)e^{b_1 x} - p(b_1)e^{b_2 x}$$

is given by

$$v_{\mathrm{eu},f}(\vartheta, x) = p(b_2)e^{b_1 x + p(b_1)\vartheta} - p(b_1)e^{b_2 x + p(b_2)\vartheta}.$$

We have $p(b_1) < 0 < p(b_2)$ and the reader easily verifies that for any $x \in \mathbb{R}$ the mapping $\vartheta \mapsto v_{\mathrm{eu},f}(\vartheta, x)$ is convex. Moreover, for any $x \in \mathbb{R}$ the infimum in the defining equation (2.2) of the embedded American option is attained at the unique point

$$\widehat{\vartheta}(x) = \min\left\{T, \frac{x}{\alpha}\mathbb{1}_{\mathbb{R}_+}\left((b_1 - b_2)\, x\right)\right\}$$

[1] If the reader does not feel comfortable with the lookahead to Proposition 3.2, the uniqueness may be obtained by recognizing that the argument proving Assertion 2 from Proposition 2.10 does not essentially depend on the continuity of the American claim.

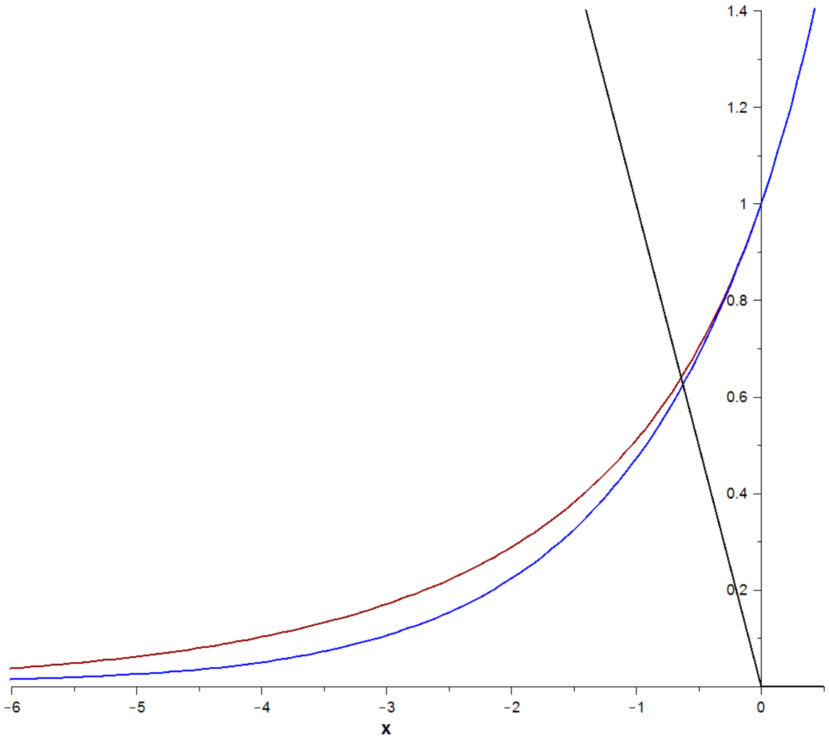

Figure 2.5: The functions f (red), $\mathrm{am}_T(f)$ (blue), $\widehat{\vartheta}$ (black) from Example 2.28 for the parameters $r = 0, \sigma = \sqrt{2}, T = 4, b_1 = \frac{1}{2}$ and $b_2 = \frac{3}{4}$.

where $\alpha := \frac{p(b_2)-p(b_1)}{b_1-b_2}$. In case that $b_1 > b_2$, the embedded American option is therefore given by

$$\mathrm{am}_T(f)(x) = \begin{cases} f(x) & \text{if } x \leq 0, \\ (p(b_2) - p(b_1)) \exp\left(\frac{b_1 p(b_2) - b_2 p(b_1)}{p(b_2)-p(b_1)} x\right) & \text{if } 0 < x < T\alpha, \\ v_{\mathrm{eu},f}(T,x) & \text{otherwise,} \end{cases}$$

and for $b_1 < b_2$ we have

$$\mathrm{am}_T(f)(x) = \begin{cases} f(x) & \text{if } x \geq 0, \\ (p(b_2) - p(b_1)) \exp\left(\frac{b_1 p(b_2) - b_2 p(b_1)}{p(b_2)-p(b_1)} x\right) & \text{if } 0 > x > T\alpha, \\ v_{\mathrm{eu},f}(T,x) & \text{otherwise.} \end{cases}$$

Proposition 2.19 warrants that $\mathrm{am}_T(f)$ is globally represented by f and that

$$C_T = \left\{(\vartheta, x) \in [0,T] \times \mathbb{R} \mid \vartheta > \widehat{\vartheta}(x)\right\}.$$

In particular, we obtain that the early exercise boundary is parametrized by the curve $\widehat{\vartheta}$. Clearly, the embedded American option is piecewise analytic on the set $\mathbb{R} \setminus \{0, T\alpha\}$. The reader may compare this observation to [JM1, Proposition 17].

An elementary calculation shows that the function $\mathrm{am}_T(f)$ is continuously differentiable at $x = 0$ and $x = T\alpha$. Moreover, the interested reader may verify that the second order derivative of $\mathrm{am}_T(f)$ does not exist at the latter points. To put differently, we have *first order but not second order smooth fit.* Figure 2.5 depicts the graphs of the functions $f, \mathrm{am}_T(f)$ and $\widehat{\vartheta}$ for the parameters $r = 0, \sigma = \sqrt{2}, T = 4, b_1 = \frac{1}{2}$ and $b_2 = \frac{3}{4}$.

Another interesting example of an embedded American payoff in the Black-Scholes setting can be found in [JM1, Section 3.1]. Next, we will examine two examples where the underlying stochastic process is a standard Brownian motion.

2.29 Example: As above, let us denote by W a standard Brownian motion with $\mathbb{P}_x(W_0 = x) = 1$ for any $x \in \mathbb{R}$. We consider the Bachelier market

$$\begin{aligned} B_t &= 1, \\ \mathrm{d}X_t &= \mathrm{d}W_t \end{aligned} \tag{2.30}$$

with some time horizon $T \in [\frac{1}{6}, \infty]$. The value function associated to the European payoff

$$f(x) = \left(x^2 - \frac{1}{2}\right)^2$$

is given by

$$v_{\mathrm{eu},f}(\vartheta, x) = 3\vartheta^2 + \left(6x^2 - 1\right)\vartheta + \left(x^2 - \frac{1}{2}\right)^2.$$

For a fixed $x \in \mathbb{R}$ we find that $\vartheta \mapsto v_{\mathrm{eu},f}(\vartheta, x)$ is a convex mapping. Moreover, it is easily seen that for any $x \in \mathbb{R}$ the infimum in (2.2) is attained at the uniquely determined maturity

$$\widehat{\vartheta}(x) = \left(\frac{1}{6} - x^2\right) \mathbb{1}_{\left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}\right)}(x).$$

Proposition 2.19 warrants that the embedded American option

$$\mathrm{am}_T(f)(x) = \begin{cases} \frac{1}{6} - 2x^4 & \text{if } -\frac{1}{\sqrt{6}} < x < \frac{1}{\sqrt{6}}, \\ f(x) & \text{otherwise}, \end{cases}$$

is globally represented by the payoff f and that the curve $\widehat{\vartheta}$ parametrizes the associated early exercise boundary. Figure 2.6 depicts the mappings $f, \mathrm{am}_T(f)$ and $\widehat{\vartheta}$. Let us remark that $\mathrm{am}_T(f)$ is piecewise analytic on $\mathbb{R} \setminus \{-1/\sqrt{6}, 1/\sqrt{6}\}$ and that we have *first order smooth fit* at $x = -1/\sqrt{6}$ and $x = 1/\sqrt{6}$. Furthermore, the second order derivative of the embedded American option does not exist at the latter points.

2.31 Example: We consider the market (2.30) with some terminal time $T \in (0, \infty]$ and the European payoff

$$f(x) = 2\mathbb{1}_{\{x \leq -1\}} + (1 - x)\mathbb{1}_{\{-1 < x < 1\}}.$$

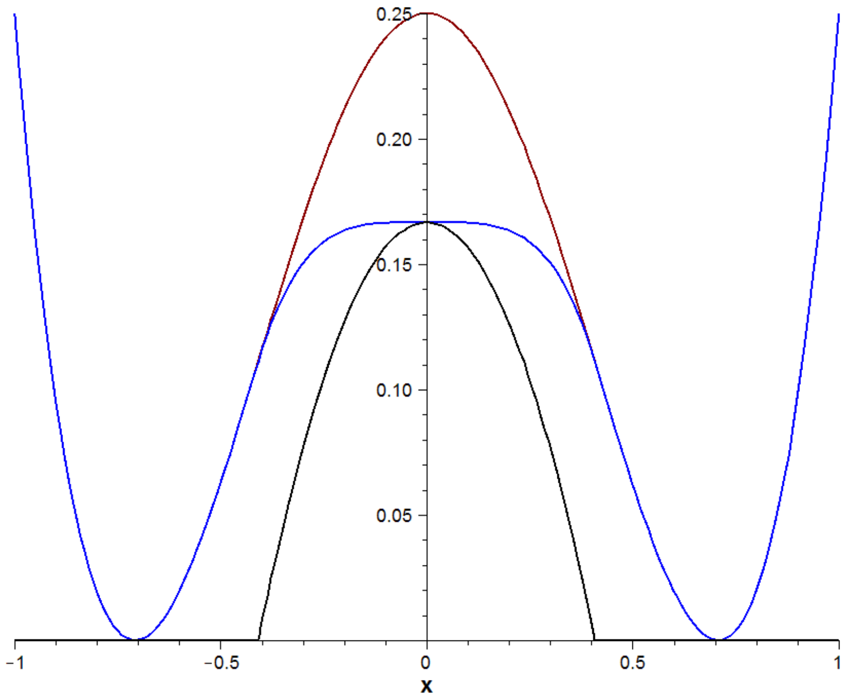

Figure 2.6: The functions f (red), $\mathrm{am}_T(f)$ (blue) and $\widehat{\vartheta}$ (black) from Example 2.29.

The associated European value function is given by

$$v_{\mathrm{eu},f}(\vartheta,x)=\frac{\sqrt{\vartheta}}{\sqrt{2\pi}}\left(e^{-\frac{(x-1)^2}{2\vartheta}}-e^{-\frac{(x+1)^2}{2\vartheta}}\right)+\operatorname{erf}\left(\frac{x-1}{\sqrt{2\vartheta}}\right)\frac{x-1}{2}-\operatorname{erf}\left(\frac{x+1}{\sqrt{2\vartheta}}\right)\frac{x+1}{2}+1$$

and therefore

$$\partial_\vartheta v_{\mathrm{eu},f}(\vartheta,x)=\frac{1}{2\sqrt{2\pi\vartheta}}\left(e^{-\frac{(x-1)^2}{2\vartheta}}-e^{-\frac{(x+1)^2}{2\vartheta}}\right).$$

The latter equation shows that the mapping $\vartheta\mapsto v_{\mathrm{eu},f}(\vartheta,x)$ is strictly decreasing for $x<0$, strictly increasing for $x>0$ and constant for $x=0$. In other words, for $x=0$ the infimum in the defining equation (2.2) of the embedded American option is attained at every $\vartheta\in[0,T]$. Hence, the embedded American option is given by

$$\mathrm{am}_T(f)(x)=\begin{cases}v_{\mathrm{eu},f}(T,x) & \text{if } x<0,\\ 1-x & \text{if } 0\le x\le 1,\\ 0 & \text{otherwise.}\end{cases}$$

Figure 2.7 depicts the graphs of the functions $f,\mathrm{am}_\infty(f)$ and $\mathrm{am}_T(f)$ for different time horizons $T\in\{10^k \mid k=-1,...,4\}$. Let H_M be defined as in (2.17). Obviously, we have $H_M(\vartheta,x)=1$ for any $(\vartheta,x)\in(0,T]\times\mathbb{R}_{++}$ such that $\vartheta<\infty$. The second assertion of Proposition 2.16 implies that the continuation and stopping set associated to the payoff $\mathrm{am}_T(f)$ are given by

$$\begin{aligned}C_T&=\{(\vartheta,x)\in(0,T]\times\mathbb{R}\mid x>0 \text{ and } \vartheta<\infty\},\\ S_T&=\{(\vartheta,x)\in(0,T]\times\mathbb{R}\mid x\le 0 \text{ and } \vartheta<\infty\}.\end{aligned}\tag{2.32}$$

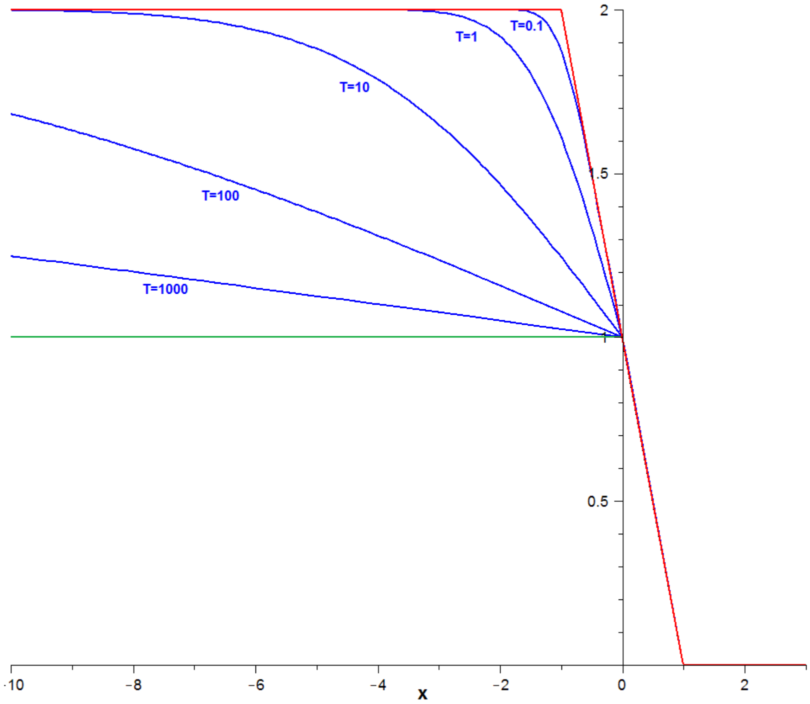

Figure 2.7: The functions f (red), $\mathrm{am}_T(f)$ (blue), $\mathrm{am}_\infty(f)$ (green) from Example 2.31.

Owing to the first assertion of the latter proposition, we find that $\mathrm{am}_T(f)$ is globally represented by f in the sense of Definition 2.8. Moreover, by virtue of Proposition 2.10, Assertion 6 we conclude that C_T is a connected component of the continuation region associated to the *American butterfly* payoff

$$g(x) := (1+x)\mathbb{1}_{[-1,0)}(x) + (1-x)\mathbb{1}_{[0,1]}(x)$$

and that the latter is represented by f on C_T. In contrast to the Examples 2.28 and 2.29 we do not have smooth fit. Indeed, the reader easily verifies that

$$\partial_x v_{\mathrm{eu},f}(T,x) = \frac{1}{2}\mathrm{erf}\left(\frac{x-1}{\sqrt{2T}}\right) - \frac{1}{2}\mathrm{erf}\left(\frac{x+1}{\sqrt{2T}}\right)$$

holds true for any $x \in \mathbb{R}$. Consequently, we obtain

$$\partial_x^- \mathrm{am}_T(f)(0) = -\mathrm{erf}\left(\frac{1}{\sqrt{2T}}\right) > -1 = \partial_x^+ \mathrm{am}_T(f)(0)$$

which shows that there is *no smooth fit* at the point $x = 0$. Here we denote by ∂_x^- and ∂_x^+ the left and right derivative, respectively.

Another interesting example of an embedded American option will be studied in Subsection 2.1.3 below. There we will discuss some results related to the EAO generated by the European put. This particular example teaches us that the representability of an American payoff may depend on the time horizon of the market.

2.1.3 The American option embedded into the European put

Differing from the previous notational conventions, we will use non-logarithmic stock prices $s \in \mathbb{R}_{++}$ until the end of this section. Whenever necessary, we ask the reader to adapt the notations $v_{\mathrm{eu},f}, v_{\mathrm{am},g}, C_T, \mathrm{am}_T(f)$ from Section 1.2 and Subsection 2.1.1. As above, the probability density function and the cumulative distribution function of a standard normal random variable are denoted by $\varphi := \mathrm{N}(0,1,\cdot)$ and Φ, respectively.

Let us consider a European put option $f(s) := (K-s)_+$ with strike $K \in \mathbb{R}_{++}$ in a risk-neutral Black-Scholes market with interest rate $r \in \mathbb{R}_{++}$ and volatility $\sigma \in \mathbb{R}_{++}$. Owing to the Black-Scholes formula, the associated European value function is given by

$$P(\vartheta, s) := v_{\mathrm{eu},f}(\vartheta, s) = e^{-r\vartheta} K \Phi(-d_2(\vartheta, s)) - s\,\Phi(-d_1(\vartheta, s)) \tag{2.33}$$

with

$$\begin{aligned} d_1(\vartheta, s) &:= \frac{\log(s/K) + (r + \sigma^2/2)\vartheta}{\sigma\sqrt{\vartheta}} \\ d_2(\vartheta, s) &:= \frac{\log(s/K) + (r - \sigma^2/2)\vartheta}{\sigma\sqrt{\vartheta}} \end{aligned}$$

where s denotes the spot price of the underlying and $\vartheta \in \mathbb{R}_+$ the maturity of the option. We will prove that for any sufficiently small terminal time T, there exists a continuous curve $\widehat{\vartheta} : \mathbb{R}_{++} \to [0,T]$ such that $\mathrm{am}_T(f)(s) = P(\widehat{\vartheta}(s), s)$ holds true for any $s \in \mathbb{R}_{++}$. Proposition 2.19 then warrants that the American payoff $\mathrm{am}_T(f)$ is globally represented by f up to time T. We define $P^{(k,l)} := \partial_\vartheta^k \partial_s^l P$ and recall the following well-known facts:

$$\begin{aligned} P^{(1,0)} &= \frac{s\sigma}{2\sqrt{\vartheta}} \varphi(d_1) - rKe^{-r\vartheta}\Phi(-d_2) \\ P^{(0,1)} &= -\Phi(-d_1) \\ P^{(1,1)} &= \frac{(r + \sigma^2/2)\,\vartheta - \log\,(s/K)}{2\vartheta^{3/2}\sigma} \varphi(d_1) \end{aligned}$$

Consequently, for any $(\vartheta, s) \in \mathbb{R}_+ \times \mathbb{R}_{++}$ we have

$$P^{(1,1)}(\vartheta, s) > 0 \tag{2.34}$$

if and only if $s < K \exp\left((r + \sigma^2/2)\vartheta\right)$. Furthermore, the reader easily verifies that for any $T > 0$ the following three properties are satisfied:

$$\liminf_{s \searrow 0} \sup_{\vartheta \in [0,T]} P^{(1,0)}(\vartheta, s) < 0 \tag{2.35}$$

$$\lim_{\vartheta \searrow 0} P^{(1,0)}(\vartheta, K) = \infty \tag{2.36}$$

$$\forall s \in (0, K) \ : \ \lim_{\vartheta \searrow 0} P^{(1,0)}(\vartheta, s) = -rK \tag{2.37}$$

Owing to (2.36), we can choose some constant $\gamma > 0$ such that $P^{(1,0)}(\vartheta, K) > 0$ holds true for any $\vartheta \in (0, \gamma)$. In the following, let T always denote some terminal time satisfying

$0 < T < \gamma$. Property (2.35) warrants that $\liminf_{s \searrow 0} P^{(1,0)}(T, s) < 0$. Due to (2.34) and the intermediate value theorem, there exists a uniquely determined constant $K_T \in (0, K)$ such that $P^{(1,0)}(T, K_T) = 0$, $P^{(1,0)}(T, s) < 0$ for $s \in (0, K_T)$ and $P^{(1,0)}(T, s) > 0$ for $s \in (K_T, K]$. Taking (2.37) into account, we conclude that

$$m(s) := \min_{\vartheta \in [0,T]} P(\vartheta, s) < P(0, s) \wedge P(T, s)$$

holds true for any $s \in (K_T, K)$. To put differently, for any $s \in (K_T, K)$ the non-empty compact set

$$M_s := \{\vartheta \in [0, T] \mid P(\vartheta, s) = m(s)\}$$

is contained in the open interval $(0, T)$. We write

$$\widehat{\vartheta}(s) := \max M_s$$

for the largest value of the set M_s. Clearly, for any $s \in (K_T, K)$ we have $P^{(1,0)}(\widehat{\vartheta}(s), s) = 0$. By eventually decreasing the bound γ, we can always achieve that $P^{(2,0)}(\widehat{\vartheta}(s), s) > 0$ holds true for any $s \in (K_T, K)$. The latter can be verified by analyzing the asymptotic behavior of the derivative $P^{(2,0)}$ as $\vartheta \to 0$. The calculation is elementary but somewhat lengthy and therefore will be omitted. Theorem 5.7 yields that the mapping $s \mapsto \widehat{\vartheta}(s)$ is analytic on some open complex domain containing the interval (K_T, K). Moreover, owing to Equation (2.34), we have

$$(\partial_s \widehat{\vartheta})(s) = -\frac{P^{(1,1)}(\widehat{\vartheta}(s), s)}{P^{(2,0)}(\widehat{\vartheta}(s), s)} < 0 \tag{2.38}$$

for any $s \in (K_T, K)$ and this clearly implies that the limits $\lim_{s \searrow K_T} \widehat{\vartheta}(s)$ and $\lim_{s \nearrow K} \widehat{\vartheta}(s)$ exist. A simple calculation shows that $\widehat{\vartheta}(K) := 0$ continuously extends the curve into the point $s = K$. Indeed, assuming $v_1 := \lim_{s \nearrow K} \widehat{\vartheta}(s) > 0$ yields $\widehat{\vartheta}(s) \in (v_1, T]$ for any $s \in (K_T, K)$. The mapping $P^{(1,0)}$ is continuous on $\mathbb{R}^2_{++}$ and consequently we obtain the contradiction $0 < P^{(1,0)}(v_1, K) = \lim_{s \nearrow K} P^{(1,0)}(\vartheta(s), s) = 0$. Furthermore, by eventually further decreasing γ, we can achieve that $\widehat{\vartheta}(K_T) := T$ extends the curve continuously into $s = K_T$.

Obviously, for any $s \geq K$ the mapping $[0, T] \ni \vartheta \mapsto P(\vartheta, s)$ attains its unique minimum at $\vartheta = 0$. Given any $s \in [K_T, K)$, we will now prove that minimum of the mapping $[0, T] \ni \vartheta \mapsto P(\vartheta, s)$ is uniquely attained at $\widehat{\vartheta}(s)$. Assuming that the latter is false, i.e. $|M_s| \geq 2$, we can find some maturity $\rho \in (0, \widehat{\vartheta}(s))$ such that $P(\rho, s) = P(\widehat{\vartheta}(s), s) = m(s)$. Necessarily, we have $P^{(1,0)}(\rho, s) = 0$ and taking (2.38) into account, we conclude that $s < \widehat{\vartheta}^{-1}(\rho)$. Property (2.34) now yields the contradiction $0 = P^{(1,0)}(\rho, s) < P^{(1,0)}(\rho, \widehat{\vartheta}^{-1}(\rho)) = 0$ and therefore $|M_s| = 1$ must indeed hold true.

A similar argument shows that for any $s \in (0, K_T)$ the minimum of the mapping $[0, T] \ni \vartheta \mapsto P(\vartheta, s)$ is uniquely attained at $\vartheta = T$. Indeed, owing to (2.37), no minimum can be located at $\vartheta = 0$. Now assume that for some $s \in (0, K_T)$ a minimum is located at some maturity $\rho \in (0, T)$. Clearly, the first order condition $P^{(1,0)}(\rho, s) = 0$ is satisfied and we have $s < \widehat{\vartheta}^{-1}(\rho)$. By virtue of (2.34) we conclude that $0 = P^{(1,0)}(\rho, s) < P^{(1,0)}(\rho, \widehat{\vartheta}^{-1}(\rho)) = 0$, which is clearly not possible. Let us summarize our results.

2.39 Lemma: *For any choice of model parameters $\sigma, r, K > 0$ there exists a constant $\gamma > 0$ such that for any time horizon $T \in (0, \gamma)$ the following holds true:*

1. *For any $s > 0$ there exists a uniquely determined maturity $\widehat{\vartheta}(s) \in [0, T]$ such that $P(\widehat{\vartheta}(s), s) = \min_{\vartheta \in [0,T]} P(\vartheta, s)$.*
2. *The curve $\mathbb{R}_{++} \ni s \mapsto \widehat{\vartheta}(s)$ and the embedded American option $\mathrm{am}_T(f)$ are continuous mappings.*
3. *There exists a uniquely determined constant $K_T \in (0, K)$ such that the curve $\widehat{\vartheta}$ is strictly decreasing on the interval (K_T, K). Moreover, we have $\widehat{\vartheta}(s) = T$ for any $s \in (0, K_T]$ and $\widehat{\vartheta}(s) = 0$ for any $s \geq K$.*
4. *The functions $\widehat{\vartheta}$ and $\mathrm{am}_T(f)$ are analytic on some open complex domain containing the set $\mathbb{R} \setminus \{K_T, K\}$.*
5. *The set*
$$\{(\vartheta, s) \in (0, T] \times (K_T, \infty) \mid \vartheta > \widehat{\vartheta}(s)\}$$
corresponds to the continuation region C_T associated to the American payoff $\mathrm{am}_T(f)$ and we have $f \xrightarrow{T} \mathrm{am}_T(f)$.

For illustrative purposes, let us consider the model parameters $r = 0.06, \sigma = 0.4$ and $K = 100$. We choose the time horizon $T = 1$ and remark that the latter satisfies the condition $T < \gamma$ from Lemma 2.39. Figure 2.8 depicts a numerical approximation to the curve $\widehat{\vartheta}$. Numerically solving the non-linear equation $P^{(1,0)}(T, s) = 0$ for the variable s yields $K_T \approx 69.296$. The values of the embedded American option $\mathrm{am}_T(f)$ on the interval $[K_T, K]$ are displayed in Figure 2.9.

The discussion from above provides us with a further insight: *The representability of an American claim may depend on the time horizon of the model.* In order to prove this assertion, suppose that $\sigma, r, K > 0$ and choose a terminal time T_0 satisfying the condition $T_0 \in (0, \gamma)$ from Lemma 2.39. The latter lemma implies that the American payoff $g := \mathrm{am}_{T_0}(f)$ is globally represented by the European put f up to time T_0 and that there exists a continuous curve $\widehat{\vartheta} : (K_{T_0}, \infty) \to [0, T_0]$ and such that

$$C_{T_0} = \{(\vartheta, s) \in (0, T_0] \times (K_{T_0}, \infty) \mid \vartheta > \widehat{\vartheta}(s)\}.$$

With regard to the Black-Scholes formula (2.33), it is apparent that $\lim_{T \to \infty} v_{\mathrm{eu},f}(T, s) = 0$ for any $s > 0$. Consequently, we can pick a time horizon $T \in (T_0, \infty)$ such that $v_{\mathrm{eu},f}(T, \alpha) < g(\alpha)$ where $\alpha := (K_{T_0} + K)/2$. Assertion 3 of Lemma 2.39 warrants that (T_0, α) is contained in C_{T_0}. Moreover, the monotonicity of the American value function $v_{\mathrm{am},g}$ in the variable ϑ yields

$$C_{T_0} \subset \{(\vartheta, s) \in (0, T] \times (K_{T_0}, \infty) \mid \vartheta > \widehat{\vartheta}(s)\} \subset C_T. \tag{2.40}$$

Now suppose that there exists some European payoff $\tilde{f}$ which represents g relative to (T, α). Equation (2.40) shows that the points (T_0, α) and (T, α) are located within the same connected component of C_T. By virtue of Proposition 2.10 we conclude that the mappings f and $\tilde{f}$ coincide up to a Lebesgue nullset and therefore $v_{\mathrm{am},g}(T,\alpha) = v_{\mathrm{eu},\tilde{f}}(T,\alpha) = v_{\mathrm{eu},f}(T,\alpha) < g(\alpha)$. This is clearly a contradiction which proves that the American payoff g is indeed only representable up to some maximal time horizon.

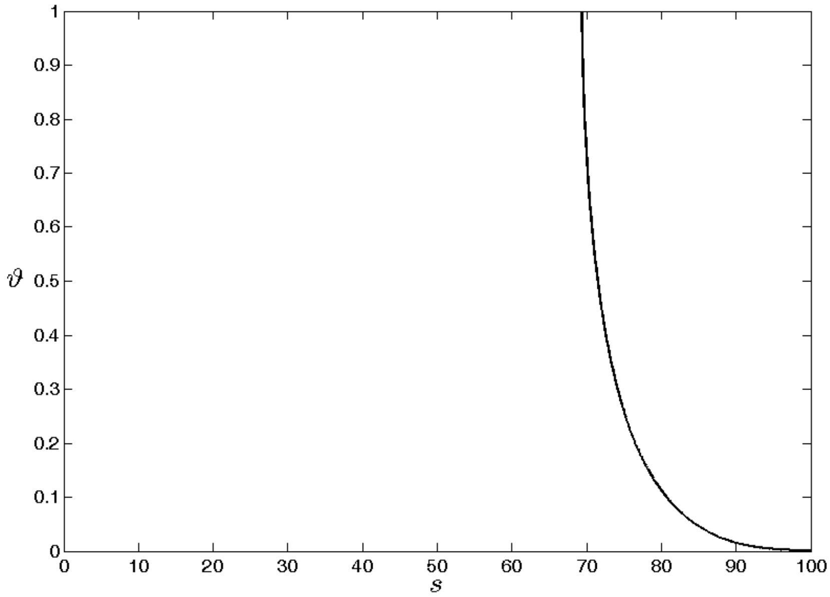

Figure 2.8: The curve $\widehat{\vartheta}$ for the parameter set $r = 0.06, \sigma = 0.4, T = 1, K = 100$.

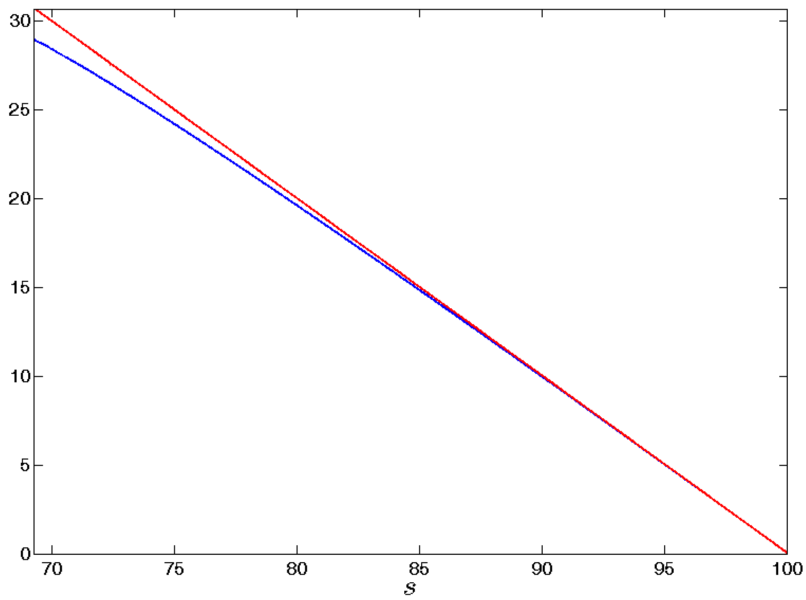

Figure 2.9: The put payoff (red) and the EAO $\mathrm{am}_T(f)$ (blue) for the parameter set $r = 0.06, \sigma = 0.4, T = 1, K = 100$.

2.2 Existence and a verification theorem

In this section we will establish the existence of cheapest dominating European options, at least in a distributional sense. Furthermore, we provide a verification theorem which warrants that an American payoff is generated and represented by its cheapest dominating European option. In order to keep things simple, we confine our theory to the Black-Scholes scenario (1.3) and consider a certain class of conveniently structured American payoffs which contains the American put. Clearly, the results below can be generalized in many aspects. Possible extensions are discussed below. Nonetheless, we deem to verify rigorously the representability of the Black-Scholes American put as the more important pending task.

Recall that our basic model consists of a deterministic bond B and a stock with log-price process X which evolve according to the stochastic differential equation

$$\begin{aligned} \mathrm{d}B_t &= rB_t\,\mathrm{d}t, \\ \mathrm{d}X_t &= \hat{r}\mathrm{d}t + \sigma\,\mathrm{d}W_t \end{aligned} \tag{2.41}$$

where $r \in \mathbb{R}_+$, $\sigma, B_0 \in \mathbb{R}_{++}, \hat{r} := r - \frac{\sigma^2}{2}$ and $\mathbb{P}_x(X_0 = x) = 1$ for any $x \in \mathbb{R}$. Moreover, assume a finite time horizon $T \in \mathbb{R}_{++}$. We focus exclusively on American payoff functions $g : \mathbb{R} \to \mathbb{R}_+$ of the type

$$g(x) = \mathbb{1}_{(-\infty,K]}(x)\ \phi(x) \tag{2.42}$$

where $K \in \mathbb{R}$ and ϕ denotes some function which is strictly positive on $(-\infty, K)$ and analytic on a complex domain containing the set $(-\infty, K)$ with $\phi(K) = 0$. Furthermore, we require that the growth condition

$$\lim_{x\to-\infty} e^{\frac{2r}{\sigma^2}x} g(x) = 0 \tag{2.43}$$

is satisfied. For example, the payoff $g(x) = \left(e^K - e^x\right)_+$ associated to the American put satisfies the latter requirements. We suppose that the necessary concavity condition

$$c(x) := g''(x) - \frac{2r}{\sigma^2}(g(x) - g'(x)) - g'(x) \leq 0, \quad x \in (-\infty, K) \tag{2.44}$$

from Equation (2.22) in Proposition 2.19 is satisfied. Note that $c(x) = -\frac{2r}{\sigma^2}e^K \leq 0$ holds true for the American put with strike price e^K. The structural requirement of (2.42) directly implies that the set $[0, T] \times (K, \infty)$ is contained in the continuation region associated to the American payoff g. Our first goal is to show that for any $x_0 > K$ the cheapest dominating European option of g relative to T, x_0 exists in a suitably generalized sense. If $f : \mathbb{R} \to \mathbb{R}_+$ denotes a European payoff function, we have

$$v_{\mathrm{eu},f}(\vartheta, x) = e^{-r\vartheta} \int_{-\infty}^{\infty} \mathrm{N}(x + \hat{r}\vartheta, \sigma^2\vartheta, y) f(y)\mathrm{d}y. \tag{2.45}$$

Here $\mathrm{N}(\mu, \sigma^2, \cdot\,)$ denotes the probability density function of the Gaussian variable $\mathcal{N}(\mu, \sigma^2)$. Put differently, we have

$$v_{\mathrm{eu},f}(\vartheta, x) = e^{-r\vartheta} \int_{-\infty}^{\infty} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \mathrm{d}\mu(y) \tag{2.46}$$

for the measure $\mu \in \mathcal{M}^+(\mathbb{R})$ with density f relative to the law of $\mathcal{N}(x_0 + \hat{r}T, \sigma^2 T)$. In the European valuation problem, the payoff function f is only needed for defining the pricing function $v_{\mathrm{eu},f}$. In view of (2.46), we can and do therefore extend the notion of a payoff "function" to include all measures $\mu \in \mathcal{M}^+(\mathbb{R})$. In line with (2.46), we define the pricing operator $v_{\mathrm{eu},\mu} : \mathbb{R}_+ \times \mathbb{R} \to [0, \infty]$ by

$$v_{\mathrm{eu},\mu}(\vartheta, x) := e^{-r\vartheta} \int_{-\infty}^{\infty} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \mathrm{d}\mu(y), \quad (\vartheta, x) \in \mathbb{R}_{++} \times \mathbb{R} \tag{2.47}$$

and

$$v_{\mathrm{eu},\mu}(0, x) := \liminf_{\substack{(\vartheta,y)\to(0,x) \\ (\vartheta,y)\in\mathbb{R}_{++}\times\mathbb{R}}} v_{\mathrm{eu},\mu}(\vartheta, x), \quad x \in \mathbb{R}. \tag{2.48}$$

In terms of our generalized domain, the linear problem (1.15) now reads as

$$\begin{array}{ll} \text{minimize} & v_{\mathrm{eu},\mu}(T, x_0) \\ \text{subject to} & \mu \in \mathcal{M}^+(\mathbb{R}), \\ & v_{\mathrm{eu},\mu}(\vartheta, x) \geq g(x) \text{ for any } (\vartheta, x) \in [0, T] \times \mathbb{R}. \end{array} \tag{2.49}$$

In line with Definition 2.1, a minimizer μ^* is called *cheapest dominating European option* (CDEO) of g relative to T, x_0. Exploiting the structural assumption (2.42), we will show that any measure μ which is admissible in (2.49) can be transformed into another admissible measure $s(\mu)$ which is concentrated on the set $(-\infty, K]$ such that $v_{\mathrm{eu},\mu}(T, x_0) = v_{\mathrm{eu},s(\mu)}(T, x_0)$. Hence, we can equivalently consider the linear program

$$\begin{array}{ll} \text{minimize} & v_{\mathrm{eu},\mu}(T, x_0) \\ \text{subject to} & \mu \in \mathcal{M}^+(-\infty, K], \\ & v_{\mathrm{eu},\mu}(\vartheta, x) \geq g(x) \text{ for any } (\vartheta, x) \in [0, T] \times (-\infty, K) \end{array} \tag{2.50}$$

in order to establish the existence of a cheapest dominating European option.

2.51 Theorem: *The optimal value of program* (2.49) *is obtained by some* $\mu^* \in \mathcal{M}^+(-\infty, K]$. *In particular, a CDEO of* g *relative to* T, x_0 *exists in the present generalized sense.*

A proof of this theorem can be found in Section 2.3. The previous result concerned the existence of the CDEO of g. We now turn to the question whether the latter generates the American claim g under consideration.

2.52 Theorem: *Let μ^* denote an optimal measure from Theorem 2.51. Suppose that for some constant $\delta > 0$ the following assumptions are satisfied:*

1. *There exists some $x_1 \in \mathbb{R}$ such that $v_{\mathrm{eu},\mu^*}(T + 2\delta, x_1) < \infty$.*
2. *For any $\vartheta \in (0, T+\delta)$ the function $x \mapsto v_{\mathrm{eu},\mu^*}(\vartheta, x) - g(x)$ assumes its unique minimum within the interval $(-\infty, K]$ at some point $\widehat{x}(\vartheta) \in (-\infty, K)$. Furthermore, we have $\liminf_{\vartheta\to 0} \widehat{x}(\vartheta) = K =: \widehat{x}(0)$.*
3. *The well-defined quantity*
$$H(\vartheta, x) := \frac{2}{\sigma^2}\partial_\vartheta v_{\mathrm{eu},\mu^*}(\vartheta, x) + \frac{2r}{\sigma^2}(v_{\mathrm{eu},\mu^*}(\vartheta, x) - g(x)) - c(x) \tag{2.53}$$
is strictly positive on the set $\{(\vartheta, \widehat{x}(\vartheta)) \mid \vartheta \in (0, T]\}$.
4. *We have $\liminf_{\vartheta\to 0} v_{\mathrm{eu},\mu^*}(\vartheta, K) < \infty$.*

Define
$$\widetilde{C}_{(T,x_0)} := \{(\vartheta, x) \in (0, T] \times \mathbb{R} \mid \ \widehat{x}(\vartheta) < x\} \tag{2.54}$$
and let $C_{(T,x_0)}$ denote the connected component of the continuation set associated to g which contains the optimization point (T, x_0), cf. Subsection 2.1.1. Then:

(a) *The function $\vartheta \to \widehat{x}(\vartheta)$ is analytic on a complex domain containing $(0, T]$.*

(b) *We have $v_{\mathrm{eu},\mu^*}(\vartheta, \widehat{x}(\vartheta)) = g(\widehat{x}(\vartheta))$ for any $\vartheta \in [0, T]$.*

(c) *The CDEO μ^* is the unique measure that* represents *g on the set $\widetilde{C}_{(T,x_0)}$ in the sense that $v_{\mathrm{am},g}(\vartheta, x) \le v_{\mathrm{eu},\mu^*}(\vartheta, x)$ for any $(\vartheta, x) \in [0, T] \times \mathbb{R}$ and equality holds on $\widetilde{C}_{(T,x_0)}$.*

(d) *The payoff g coincides on $\mathrm{cl}\,\pi(\widetilde{C}_{(T,x_0)}) = [\min_{\vartheta\in(0,T]} \widehat{x}(\vartheta), \infty)$ with the* embedded American option *of μ^* up to T in the sense that*
$$g(x) = \inf_{\vartheta\in[0,T]} v_{\mathrm{eu},\mu^*}(\vartheta, x) =: \mathrm{am}_T(\mu^*)(x) \qquad \text{for } \ x \in \pi(\widetilde{C}_{(T,x_0)}).$$

(e) *The mapping $\vartheta \mapsto \widehat{x}(\vartheta)$ is strictly increasing and we have*
$$\widetilde{C}_{(T,x_0)} = C_{(T,x_0)}.$$
Consequently, the curve $\widehat{x}$ parametrizes the early exercise boundary associated to $C_{(T,x_0)}$ and the stopping time
$$\tau_\vartheta := \inf\{t \in [0, \vartheta] \mid X_t \le \widehat{x}(\vartheta - t)\} \wedge \vartheta \tag{2.55}$$
is optimal, i.e. $v_{\mathrm{am},g}(\vartheta, x) = \mathbb{E}_x[e^{-r\tau_\vartheta} g(X_{\tau_\vartheta})]$ holds for any $(\vartheta, x) \in [0, T] \times \mathbb{R}$.

Remark: On closer inspection the proof of the theorem shows that the analyticity of g *is actually only required on a complex domain containing the interval* $[\min_{\vartheta\in(0,T]}\widehat{x}(\vartheta), K)$.

A proof of this theorem will be provided in Section 2.4. Moreover, we ask the interested reader to compare the latter theorem to the assertions of Proposition 2.10 and Proposition 2.19.

Loosely speaking, the assumptions from above serve the following purposes: Assumption 1 warrants that the generalized European value function v_{eu,μ^*} is analytic on a $\mathbb{C}^2$-domain containing the set $(0, T+2\delta)\times\mathbb{R}$. The second assumption allows us to deduce that the mapping $[0,T]\ni\vartheta\mapsto\widehat{x}(\vartheta)$ is continuous. Taking Assumption 3 into account, we can conclude that the curve $\widehat{x}$ is analytic on an open complex domain containing the interval $(0,T]$. Assumption 4 ensures that the optional sampling theorem can be applied in order to prove that $v_{\mathrm{am},g}(\vartheta,x)\geq v_{\mathrm{eu},\mu^*}(\vartheta,x)$ holds true for any $(\vartheta,x)\in\widetilde{C}_{(T,x_0)}$.

What are the strengths and weaknesses of this result? The assumptions imposed by Theorem 2.52 are in some way related to certain qualitative properties of the cheapest dominating European option. On the negative side, this means that it does not warrant representability yet unless we are able to rigorously derive these properties for the CDEO associated to the specific claim under consideration. This, however, is complicated by the fact that the CDEO is typically not known explicitly. And yet, numerically the CDEO is obtained quite easily as it is explained in Chapter 3. While the numerical approximations as such cannot tell whether the CDEO represents the American claim or just provides a relatively close upper bound, they should give a good and relatively reliable indication whether the *qualitative* properties needed for Theorem 2.52 hold true. As an illustration, we study the prime example of the American put in Section 3.2.

2.3 Existence - proof of Theorem 2.51

Let g denote an American payoff function satisfying all the requirements from Section 2.2. First, we verify that in search of a solution of Program (2.49) it is only necessary to consider measures $\mu \in \mathcal{M}^+(-\infty, K]$. To this end we define by

$$\begin{aligned}
&\mathcal{M}^+(\mathbb{R}) \ni \mu \mapsto s(\mu) := \nu_1 + \nu_2 \\
&\mathrm{d}\nu_1 := \mathbb{1}_{(-\infty,K]}\,\mathrm{d}\mu \\
&\mathrm{d}\nu_2 := \mu((K,\infty))\,\mathrm{d}\delta_K
\end{aligned}$$

the mapping which relocates any mass that is contained in (K, ∞) to the point K. Here we denote by δ_K the Dirac measure concentrated at K. The reader easily verifies that s maps onto the cone $\mathcal{M}^+(-\infty, K]$ and preserves the total variation of any non-negative measure, i.e. $\|s(\mu)\| = \|\mu\|$. Now suppose that $\mu \in \mathcal{M}^+(\mathbb{R})$ is admissible in program (2.49). Clearly, we have $v_{\mathrm{eu},\mu}(T, x_0) = e^{-rT}\|\mu\| = e^{-rT}\|s(\mu)\| = v_{\mathrm{eu},s(\mu)}(T, x_0)$. Due to the second assertion of Lemma 5.2, we can pick for any $(\vartheta, x) \in (0, T) \times (-\infty, K)$ some constant $c(\vartheta, x) > 0$ such that

$$\frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} = c(\vartheta, x) \exp\left(-\frac{(y - A(\vartheta, x))^2}{2B(\vartheta)}\right)$$

where $A(\vartheta, x) := x_0 + (x - x_0)T/(T - \vartheta)$ and $B(\vartheta) := \sigma^2 T\vartheta/(T - \vartheta)$. Recalling that $x < K < x_0$, we obtain $A(\vartheta, x) := x_0 + (x - x_0)T/(T - \vartheta) < K$ and therefore

$$\begin{aligned}
v_{\mathrm{eu},s(\mu)}(\vartheta, x) &= v_{\mathrm{eu},\nu_1}(\vartheta, x) + \mu((K,\infty))\, v_{\mathrm{eu},\delta_K}(\vartheta, x) \\
&= v_{\mathrm{eu},\nu_1}(\vartheta, x) + e^{-r\vartheta} \int_{(K,\infty)} c(\vartheta, x) \exp\left(-\frac{(K - A(\vartheta, x))^2}{2B(\vartheta)}\right) \mathrm{d}\mu(y) \\
&\geq v_{\mathrm{eu},\nu_1}(\vartheta, x) + e^{-r\vartheta} \int_{(K,\infty)} c(\vartheta, x) \exp\left(-\frac{(y - A(\vartheta, x))^2}{2B(\vartheta)}\right) \mathrm{d}\mu(y) \\
&= v_{\mathrm{eu},\nu_1}(\vartheta, x) + v_{\mathrm{eu},\mu-\nu_1}(\vartheta, x) \\
&= v_{\mathrm{eu},\mu}(\vartheta, x) \geq g(x)
\end{aligned}$$

holds true for any $(\vartheta, x) \in (0, T) \times (-\infty, K)$. Along the same lines we can apply the third assertion of Lemma 5.2 in order to obtain

$$\begin{aligned}
v_{\mathrm{eu},s(\mu)}(T, x) &= v_{\mathrm{eu},\nu_1}(T, x) + \mu((K,\infty))\, v_{\mathrm{eu},\delta_K}(T, x) \\
&\geq v_{\mathrm{eu},\nu_1}(T, x) + v_{\mathrm{eu},\mu-\nu_1}(T, x) \\
&= v_{\mathrm{eu},\mu}(T, x) \geq g(x)
\end{aligned}$$

for any $x < K$. Summing up, the calculations from above imply that the inequality $v_{\mathrm{eu},s(\mu)} \geq g$ is satisfied on the set $(0, T] \times (-\infty, K)$. The payoff g is assumed to vanish on $[K, \infty)$ and therefore the measure $s(\mu)$ is admissible in Program (2.49). In conclusion, if the optimal value in (2.49) is obtained, we can find an optimizer which is concentrated on the set $(-\infty, K]$. Hence, the linear programs (2.49) and (2.50) are indeed equivalent.

2.3.1 Transformation

Now we transform our model (2.41) to a market with constant bond price process, following the approach explained in [LU]. To this end, let

$$\begin{aligned} \widetilde{B}_t &= 1, \\ \mathrm{d}\widetilde{X}_t &= \tilde{r}\mathrm{d}t + \sigma \mathrm{d}W_t \end{aligned} \tag{2.56}$$

with $\tilde{r} := -r - \sigma^2/2 < 0$ and

$$\widetilde{g}(x) := e^{(2r/\sigma^2)x} g(x) \tag{2.57}$$

where W denotes a standard Brownian motion and $\widetilde{X}_0 = x$ holds almost surely under the measure $\mathbb{P}_x$. The growth condition (2.43) warrants that $\widetilde{g}$ is a continuous function vanishing at infinity. Invoking a measure change with density process

$$\left(e^{-(2r/\sigma^2)(X_t - X_0)} e^{-rt}\right)_{t\geq 0},$$

it is easy to see that

$$\mathbb{E}_x\left[e^{-r\tau} g(X_\tau)\right] = e^{-(2r/\sigma^2)x}\, \mathbb{E}_x\left[e^{(2r/\sigma^2)\widetilde{X}_\tau} g(\widetilde{X}_\tau)\right] = e^{-(2r/\sigma^2)x}\, \mathbb{E}_x\left[\widetilde{g}(\widetilde{X}_\tau)\right]$$

holds true for any stopping time τ. Likewise, we have $\mathbb{E}_x[e^{-r\vartheta} f(X_\vartheta)] = e^{-(2r/\sigma^2)x}\, \mathbb{E}_x[\widetilde{f}(\widetilde{X}_\vartheta)]$ for any European payoff function $f : \mathbb{R} \to \mathbb{R}_+$ and any $\vartheta \in \mathbb{R}_+$ where $\widetilde{f}(x) := e^{(2r/\sigma^2)x} f(x)$. Some simple algebraic manipulations show that

$$\begin{aligned} v_{\mathrm{eu},\mu}(\vartheta, x) &= e^{-r\vartheta} \int_{-\infty}^{K} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \mathrm{d}\mu(y) \\ &= e^{(-2r/\sigma^2)x} \int_{-\infty}^{K} \frac{\mathrm{N}\left(x + \tilde{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)} e^{(2r/\sigma^2)x_0 - rT} \mathrm{d}\mu(y). \end{aligned} \tag{2.58}$$

Consequently, the linear program (2.50) is, up to renormalizing the target functional, equivalent to

$$\begin{array}{ll} \text{minimize} & \|\mu\| \\ \text{subject to} & \mu \in \mathcal{M}^+(-\infty, K], \\ & \widetilde{v}_{\mathrm{eu},\mu}(\vartheta, x) \geq \widetilde{g}(x) \text{ for any } (\vartheta, x) \in [0,T] \times (-\infty, K) \end{array} \tag{2.59}$$

where

$$\begin{aligned} \widetilde{v}_{\mathrm{eu},\mu}(\vartheta, x) &:= \int_{-\infty}^{K} \frac{\mathrm{N}\left(x + \tilde{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)} \mathrm{d}\mu(y), \qquad && (\vartheta, x) \in \mathbb{R}_{++} \times \mathbb{R}, \\ \widetilde{v}_{\mathrm{eu},\mu}(0, x) &:= \liminf_{\substack{(\vartheta,y)\to(0,x) \\ (\vartheta,y)\in\mathbb{R}_{++}\times\mathbb{R}}} \widetilde{v}_{\mathrm{eu},\mu}(\vartheta, x), && x \in \mathbb{R}. \end{aligned}$$

Here we denote by $\|\mu\|$ the total variation norm of the measure μ. Clearly, we have $\widetilde{v}_{\mathrm{eu},\mu}(T, x_0) = \|\mu\| < \infty$ for any $\mu \in \mathcal{M}^+(\mathbb{R})$. Until the end of Section 2.3 we will continue to work in this transformed setting.

2.3.2 Duality

We define the set $\Omega := (0, T) \times (-\infty, K]$ and the linear operators

$$\mathbf{T} : \mathcal{M}(-\infty, K] \to C(\Omega) \qquad\qquad \mathbf{T}\mu(t,x) := \int_{-\infty}^{K} \kappa(t,x,y)\ \mathrm{d}\mu(y),$$

$$\mathbf{T}' : \mathcal{M}(\Omega) \to \mathcal{B}((-\infty, K], \mathbb{R}) \qquad\qquad \mathbf{T}'\lambda(y) := \int_{\Omega} \kappa(t,x,y)\ \mathrm{d}\lambda(t,x)$$

with the integral kernel

$$\begin{aligned}\kappa(t,x,y) &:= \frac{\mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)} \\ &= \sqrt{\frac{T}{T-t}} \exp\left(-\frac{(y - A(x,t))^2}{2B(t)}\right) \exp\left(\frac{(x - x_0 - \tilde{r}t)^2}{2\sigma^2 t}\right) \\ &= \frac{T}{t} \frac{\mathrm{N}(A(x,t), B(t), y)}{\mathrm{N}(x_0 + \tilde{r}t, \sigma^2 t, x)}\end{aligned} \tag{2.60}$$

where $A(t,x) := x_0 + (x - x_0)T/t$ and $B(t) := \sigma^2 T(T-t)/t$, cf. Lemma 5.2. Taking the specific structure of the integral kernel κ into account, we can show that for any measure $\mu \in \mathcal{M}(-\infty, K]$ the mapping $\Omega \ni (t,x) \mapsto \mathbf{T}\mu(t,x)$ is analytic on the open $\mathbb{C}^2$-domain

$$G := \left\{\vartheta \in \mathbb{C} \;\middle|\; \sqrt{(\operatorname{Re}\vartheta - T/2)^2 + (\operatorname{Im}\vartheta)^2} < T/2\right\} \times \mathbb{C}.$$

This is a special case of *Step 1* from Section 2.4 below and therefore we postpone a proof until then. In particular, we find that range of the operator $\mathbf{T}$ is indeed contained in $C(\Omega)$. Clearly, the definition of $\mathbf{T}$ naturally extends to $\mathcal{M}(\mathbb{R})$.

2.61 Lemma: *If $\mathbf{T}\mu = 0$ on some open subset of Ω then $\mu = 0$. In particular, we find that the operator $\mathbf{T}$ is injective.*

Proof. Let $\mu \in \mathcal{M}(\mathbb{R})$ be a measure such that $\mathbf{T}\mu$ vanishes on some open subset of Ω. Denote by $\mu = \mu^+ - \mu^-$ the Hahn-Jordan decomposition of μ. By analyticity and the identity theorem we can conclude that $\mathbf{T}\mu = 0$ holds true on $(0,T) \times \mathbb{R}$. Due to the fact that $A(x + x_0/2, T/2) = 2x$ and $B(T/2) = \sigma^2 T$, we obtain from Equation (2.60) that

$$\mathrm{N}\left(\frac{x_0 + \tilde{r}T}{2}, \sigma^2 \frac{T}{2}, x\right) \mathbf{T}\mu^\pm\left(\frac{T}{2}, x + \frac{x_0}{2}\right) = 2\int_{-\infty}^{\infty} \mathrm{N}\left(2x, \sigma^2 T, y\right) \mathrm{d}\mu^\pm(y).$$

From above we know that $\mathbf{T}\mu^+ = \mathbf{T}\mu^-$ and consequently

$$\int_{-\infty}^{\infty} \mathrm{N}\left(y, \sigma^2 T, x\right) \mathrm{d}\mu^+(y) = \int_{-\infty}^{\infty} \mathrm{N}\left(y, \sigma^2 T, x\right) \mathrm{d}\mu^-(y)$$

holds true for any $x \in \mathbb{R}$. Multiplying both sides of the latter equation with $e^{\mathrm{i}zx}$ and integrating the x-coordinate over the real line yields

$$\int_{-\infty}^{\infty} \exp\left(\mathrm{i}yz - \frac{\sigma^2 T}{2} z^2\right) \mathrm{d}\mu^+(y) = \int_{-\infty}^{\infty} \exp\left(\mathrm{i}yz - \frac{\sigma^2 T}{2} z^2\right) \mathrm{d}\mu^-(y)$$

for all $z \in \mathbb{R}$. By the injectivity of the Fourier transform on $\mathcal{M}^+(\mathbb{R})$ we conclude that the orthogonal measures μ^- and μ^+ coincide. This is only possible if $\mu = \mu^\pm = 0$ which shows that $\mathbf{T}$ is indeed injective. □

After these preliminary remarks we now return to our optimization problem. The convex program (2.59) from above can be rephrased in functional analytic terms as

$$\begin{aligned} \text{minimize} \quad & \|\mu\| \\ \text{subject to} \quad & \mathbf{T}\mu - \widetilde{g} \in C^+(\Omega), \\ & \mu \in \mathcal{M}^+(-\infty, K]. \end{aligned} \tag{P_0}$$

The requirement that the European value function dominates the payoff is expressed by the conic constraint. Please, note that we silently switched to forward time t in favor of notational convenience. To this primal minimization problem we associate the Lagrange dual

$$\begin{aligned} \text{maximize} \quad & \langle \widetilde{g}, \lambda \rangle \\ \text{subject to} \quad & \mathbf{T}'\lambda(y) \le 1 \qquad \forall y \in (-\infty, K], \\ & \lambda \in \mathcal{M}^+(\Omega) \end{aligned} \tag{D_0}$$

where

$$\langle \widetilde{g}, \lambda \rangle := \int_\Omega \widetilde{g}(x) \, \mathrm{d}\lambda(t, x).$$

Let us remark that this dual problem allows for a probabilistic or physical interpretation. To this end, suppose that particles move in space-time $\Omega \subset \mathbb{R}_+ \times \mathbb{R}$, where the first coordinate of (t, x) stands for time and the second for the location at this time. In the space coordinate x the particles are assumed to follow a Brownian motion with drift $\tilde{r}$ and diffusion coefficient σ. Let us inject particles of total mass $\lambda(\Omega)$ into Ω, distributed according to λ, i.e. mass $\lambda(A)$ is assigned to any set $A \in \mathcal{B}(\Omega)$. Where in $\mathbb{R}$ are the particles to be found at the end time T? Since they follow Brownian motion, they are distributed according to the Lebesgue density

$$y \mapsto \int_\Omega \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}\lambda(t, x).$$

On the other hand, the constraint

$$\int_\Omega \kappa(t, x, y) \, \mathrm{d}\lambda(t, x) \le 1$$

can be rephrased as

$$\int_\Omega \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}\lambda(t, x) \le \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right). \tag{2.62}$$

The right-hand side is the probability density function at time T of a Brownian motion started in x_0 at time 0. Put differently, the constraint (2.62) means that we consider only laws λ on space-time Ω such that the resulting final distribution on $\mathbb{R}$ is dominated by the Gaussian law stemming from a Brownian motion started in x_0 at time 0.

Regarding the primal problem P_0 and its formal dual D_0, we may wonder whether weak or even strong duality holds, optimizers exist and, if this is the case, whether the optimizers are linked by some complementary slackness condition. The following first main result shows that this is indeed the case, at least if the CDEO payoff strictly dominates the American payoff function $\widetilde{g}$ at all $x < K$.

2.63 Theorem:

1. *The optimal value of P_0 is obtained by some $\mu_0 \in \mathcal{M}^+(-\infty, K]$ and coincides with the optimal value of D_0. The measure μ_0 puts mass on every open subset of $(-\infty, K)$. In particular, the CDEO of $\widetilde{g}$ exists in the present generalized sense.*

2. *If $\widetilde{v}_{\mathrm{eu},\mu_0}(0, x) > \widetilde{g}(x)$ holds true for all $x \in (-\infty, K)$, the optimal value of D_0 is obtained by some measure $\lambda_0 \in \mathcal{M}^+(\Omega)$. In this case the following complementary slackness conditions are satisfied:*

$$\begin{aligned} \mathbf{T}\mu_0(\vartheta, x) &= \widetilde{g}(x) \quad \lambda_0\text{-}a.e.\ on\ \Omega, && (2.64)\\ \mathbf{T}'\lambda_0(x) &= 1 \qquad \mu_0\text{-}a.e.\ on\ (-\infty, K]. && (2.65) \end{aligned}$$

In light of the discussion from Subsection 2.3.1, the latter theorem can be easily restated in terms of the untransformed quantities $v_{\mathrm{eu},\mu}$ and g associated to the program (2.50). We immediately obtain Theorem 2.51 from the first assertion of Theorem 2.63.

2.3.3 Proof of Theorem 2.63

Let Ω and $\mathbf{T}$ be defined as above. For any $\varepsilon \in (0, T)$ we define the set

$$\Omega_\varepsilon := [\varepsilon, T - \varepsilon] \times [-1/\varepsilon, K]$$

and the following linear operator:

$$\mathbf{T}^* : \mathcal{M}(\Omega_\varepsilon) \to C_0(-\infty, K] \qquad\qquad \mathbf{T}^*\lambda(y) := \int_{\Omega_\varepsilon} \kappa(t, x, y)\, \mathrm{d}\lambda(t, x)$$

The range of the operator $\mathbf{T}^*$ is contained in $C_0(-\infty, K]$ due to Lebesgue's dominated convergence theorem, formula (2.60) and the compactness of the set Ω_ε. On the Cartesian products $C(\Omega_\varepsilon) \times \mathcal{M}(\Omega_\varepsilon)$ and $C_0(-\infty, K] \times \mathcal{M}(-\infty, K]$ we consider the algebraic pairing

$$\langle f, \nu \rangle \mapsto \int f \mathrm{d}\nu. \tag{2.66}$$

The reader verifies that the latter mapping is finitely valued, bilinear and point separating. We equip $C(\Omega_\varepsilon), \mathcal{M}(\Omega_\varepsilon)$ and $\mathcal{M}(-\infty, K]$ with the weak topologies $\sigma(C, \mathcal{M}), \sigma(\mathcal{M}, C)$ and $\sigma(\mathcal{M}, C_0)$ induced by (2.66). The function space $C_0(-\infty, K]$ is endowed with the topology of uniform convergence $\mathcal{T}_{\mathrm{uc}}$. This procedure turns all four spaces into locally convex Hausdorff spaces. For a very brief introduction to locally convex spaces we refer the reader to Section 5.3. Moreover, each space of measures is the continuous dual of the associated function space and vice versa, cf. Lemma 5.15 and [RD2, Theorem 6.19]. Fubini's theorem shows that for all measures $\mu \in \mathcal{M}(-\infty, K]$ and $\lambda \in \mathcal{M}(\Omega_\varepsilon)$ the relation

$$\langle \mathbf{T}\mu, \lambda \rangle = \langle \mu, \mathbf{T}^*\lambda \rangle \tag{2.67}$$

holds true. By virtue of Lemma 5.17 we find that the operator $\mathbf{T}$ is $\sigma(\mathcal{M}, C_0)$-$\sigma(C, \mathcal{M})$ continuous and that $\mathbf{T}^*$ is $\sigma(\mathcal{M}, C)$-$\sigma(C_0, \mathcal{M})$ continuous. Figure 2.10 below summarizes the setting.

$$\begin{array}{ccc}
\left\langle \begin{array}{c} \mathcal{M}(-\infty, K] \\ \sigma(\mathcal{M}, C_0) \end{array} \right. & , & \left. \begin{array}{c} C_0(-\infty, K] \\ \mathcal{T}_{\mathrm{uc}} \end{array} \right\rangle \\
\mathbf{T} \Big\downarrow & & \Big\uparrow \mathbf{T}^* \\
\left\langle \begin{array}{c} C(\Omega_\varepsilon) \\ \sigma(C, \mathcal{M}) \end{array} \right. & , & \left. \begin{array}{c} \mathcal{M}(\Omega_\varepsilon) \\ \sigma(\mathcal{M}, C) \end{array} \right\rangle
\end{array}$$

Figure 2.10: The paired spaces occurring in the proof of Theorem 2.63.

We want to find a measure $\mu_0 \in \mathcal{M}^+(-\infty, K]$ which solves the linear program P_0 from Subsection 2.3.2. Our strategy is to approximate the latter optimization problem by the following sequence of linear programs with milder constraints

$$\begin{aligned} &\text{minimize} && \|\mu\| \\ &\text{subject to} && \mathbf{T}\mu - \tilde{g} \in C^+(\Omega_\varepsilon), \\ &&& \mu \in \mathcal{M}^+(-\infty, K]. \end{aligned} \tag{P_ε}$$

The solution of P_0 will be obtained by compactness from the family of P_ε-extremal elements. For each $\varepsilon \in (0, \frac{T}{2})$, the Lagrange dual problem associated to P_ε is given by

$$\begin{aligned} &\text{maximize} && \langle \tilde{g}, \lambda \rangle \\ &\text{subject to} && 1 - \mathbf{T}^*\lambda \in C^+(-\infty, K], \\ &&& \lambda \in \mathcal{M}^+(\Omega_\varepsilon). \end{aligned} \tag{D_ε}$$

The optimal values of P_ε and D_ε are denoted by p_ε and d_ε, respectively. By construction we find that **weak duality** $0 \leq d_\varepsilon \leq p_\varepsilon$ holds. Indeed, in virtue of the adjointness relation (2.67) we obtain

$$0 \leq \langle \tilde{g}, \lambda \rangle \leq \langle \mathbf{T}\mu, \lambda \rangle = \langle \mu, \mathbf{T}^*\lambda \rangle \leq \langle \mu, 1 \rangle = \|\mu\| \tag{2.68}$$

for any primal admissible $\mu \in \mathcal{M}^+(-\infty, K]$ and any dual admissible $\lambda \in \mathcal{M}^+(\Omega_\varepsilon)$. Next, we verify **primal and dual attainment**. The non-negative measure with Lebesgue density

$$\frac{\mathrm{d}\tilde{\mu}}{\mathrm{d}y} := 2\|\tilde{g}\|_\infty \, \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right) \mathbb{1}_{(-\infty, K)}(y)$$

is P_ε-admissible because for any $(t,x) \in \Omega_\varepsilon$ we have

$$\begin{aligned} \mathbf{T}\tilde{\mu}(t,x) &= 2\|\tilde{g}\|_\infty \int_{-\infty}^{K} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \mathrm{d}y \\ &= 2\|\tilde{g}\|_\infty \, \mathbb{P}\left(\mathcal{N}(0,1) \leq \frac{K-x}{\sigma\sqrt{T-t}} - \tilde{r}\frac{\sqrt{T-t}}{\sigma}\right) \\ &\geq 2\|\tilde{g}\|_\infty \, \mathbb{P}\left(\mathcal{N}(0,1) \leq 0\right) = \|\tilde{g}\|_\infty. \end{aligned} \tag{2.69}$$

Obviously, the total mass of the measure $\tilde{\mu}$ is bounded by the constant $2\|\tilde{g}\|_\infty$. Therefore solving the minimization problem P_ε is equivalent to minimizing the total variation norm over the $\sigma(\mathcal{M}, C_0)$-compact set

$$C_p^\varepsilon \;:=\; \mathbf{T}^{-1}\left(\;\tilde{g} + C^+(\Omega_\varepsilon)\;\right) \;\cap\; \mathcal{M}^+(-\infty, K] \;\cap\; B_{\mathcal{M}(\mathbb{R})}(2\|\tilde{g}\|_\infty). \tag{2.70}$$

The $\sigma(\mathcal{M}, C_0)$-compactness of C_p^ε is easily established: First we note that the set $\tilde{g} + C^+(\Omega_\varepsilon)$ is homeomorphic to the $\sigma(C, \mathcal{M})$-closed cone

$$C^+(\Omega_\varepsilon) = \cap_{\lambda \in \mathcal{M}^+(\Omega_\varepsilon)} \{f \in C(\Omega_\varepsilon) \mid \langle f, \lambda \rangle \geq 0\}$$

and the continuity properties of the operator $\mathbf{T}$ warrant the $\sigma(\mathcal{M}, C_0)$-closedness of the preimage $\mathbf{T}^{-1}\left(\,\tilde{g} + C^+(\Omega_\varepsilon)\,\right)$. Secondly, we observe that the cone

$$\mathcal{M}^+(-\infty, K] = \cap_{f \in C_0^+(-\infty,K]} \{\mu \in \mathcal{M}(-\infty, K] \mid \langle f, \mu \rangle \geq 0\}$$

is $\sigma(\mathcal{M}, C_0)$-closed as well and that $B_{\mathcal{M}(\mathbb{R})}(2\|\tilde{g}\|_\infty)$ is a $\sigma(\mathcal{M}, C_0)$-compact set due to Theorem 5.18. The target functional $\mu \mapsto \|\mu\|$ is lower semi-continuous with respect to the topology $\sigma(\mathcal{M}, C_0)$ and therefore its minimal value p_ε is attained by some measure $\mu_\varepsilon \in C_p^\varepsilon$, cf. Lemma 5.9.
Next, we prove the attainment of the D_ε-optimal value. For every measure $\lambda \in \mathcal{M}(\Omega_\varepsilon)$ and $y \in (-\infty, K]$ we define $\mathbf{U}\lambda(y) := \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right) \mathbf{T}^*\lambda(y)$. Obviously $\mathbf{U}$ is a $\sigma(\mathcal{M}, C)$-$\sigma(C_0, \mathcal{M})$-continuous, linear operator from $\mathcal{M}(\Omega_\varepsilon)$ into the space $C_0(-\infty, K]$. The inequality constraint of the program D_ε is equivalent to

$$\mathbf{U}\lambda(y) \leq \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)$$

for all $y \in (-\infty, K]$. Integrating this inequality over the interval $(-\infty, K]$ yields:

$$\int_{\Omega_\varepsilon} \int_{-\infty}^{K} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right)\; \mathrm{d}y \, \mathrm{d}\lambda(t,x) \leq \int_{-\infty}^{K} \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right) \mathrm{d}y \leq 1$$

A calculation similar to (2.69) yields

$$\int_{-\infty}^{K} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \mathrm{d}y \geq \mathbb{P}\left(\mathcal{N}(0,1) \leq 0\right) = \frac{1}{2}$$

for any $(t,x) \in \Omega_\varepsilon$ and consequently any D_ε-admissible measure λ must satisfy $\|\lambda\| \leq 2$. Solving the maximization problem D_ε is therefore equivalent to maximizing the $\sigma(\mathcal{M}, C)$-continuous mapping $\lambda \mapsto \langle \tilde{g}, \lambda \rangle$ over the set

$$C_d^\varepsilon \;:=\; \mathbf{U}^{-1}\left(\;\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, \cdot\right) - C_0^+(-\infty, K]\;\right) \;\cap\; \mathcal{M}^+(\Omega_\varepsilon) \;\cap\; B_{\mathcal{M}(\Omega_\varepsilon)}(2).$$

The reader can easily modify the arguments succeeding Equation (2.70) from above in order to verify that C_d^ε is a $\sigma(\mathcal{M}, C)$-compact subset of $\mathcal{M}(\Omega_\varepsilon)$. Hence, the target functional of the Lagrange dual D_ε attains its maximal value d_ε at some measure $\lambda_\varepsilon \in C_d^\varepsilon$.

In order to prove **strong duality** $d_\varepsilon = p_\varepsilon$, we utilize different well-established techniques from convex optimization. Due to the fact that the structure of our problem is standard, we decided to keep the amount of technical prerequisites to a minimum. That is to say, we prefer elementary calculations over heavy machinery from optimization theory. We refer the reader to [RO] for a nicely written introduction to conjugate duality and optimization on paired spaces. A short summary of the most important notions is provided in Section 5.4. The Lagrange function $K : \mathcal{M}(-\infty, K] \times \mathcal{M}(\Omega_\varepsilon) \to [-\infty, \infty]$ associated to the P_ε-D_ε-duality is defined by

$$K(\mu, \lambda) := \|\mu\| + \langle \widetilde{g}, \lambda\rangle - \langle \mathbf{T}\mu, \lambda\rangle + \mathcal{I}_{\mathcal{M}^+(-\infty,K]}(\mu) - \mathcal{I}_{\mathcal{M}^+(\Omega_\varepsilon)}(\lambda) \tag{2.71}$$

where

$$\mathcal{I}_M(x) := \begin{cases} 0 & \text{if } x \in M, \\ \infty & \text{if } x \notin M \end{cases}$$

for any set M. For later reference we provide the following explicit calculations:

$$\begin{aligned} \sup_{\lambda\in\mathcal{M}(\Omega_\varepsilon)} \inf_{\mu\in\mathcal{M}(-\infty,K]} K(\mu,\lambda) &= \sup_{\lambda\in\mathcal{M}^+(\Omega_\varepsilon)} \inf_{\mu\in\mathcal{M}^+(-\infty,K]} \{\|\mu\| + \langle \widetilde{g} - \mathbf{T}\mu, \lambda\rangle\} \\ &= \sup_{\lambda\in\mathcal{M}^+(\Omega_\varepsilon)} \left\{ \langle \widetilde{g}, \lambda\rangle + \inf_{\mu\in\mathcal{M}^+(-\infty,K]} \langle 1 - \mathbf{T}^*\lambda, \mu\rangle \right\} \\ &= \sup_{\substack{\lambda\in\mathcal{M}^+(\Omega_\varepsilon) \\ \mathbf{T}^*\lambda\le 1}} \langle \widetilde{g}, \lambda\rangle = d_\varepsilon \end{aligned} \tag{2.72}$$

$$\begin{aligned} \inf_{\mu\in\mathcal{M}(-\infty,K]} \sup_{\lambda\in\mathcal{M}(\Omega_\varepsilon)} K(\mu,\lambda) &= \inf_{\mu\in\mathcal{M}^+(-\infty,K]} \sup_{\lambda\in\mathcal{M}^+(\Omega_\varepsilon)} \{\|\mu\| + \langle \widetilde{g} - \mathbf{T}\mu, \lambda\rangle\} \\ &= \inf_{\mu\in\mathcal{M}^+(-\infty,K]} \left\{ \|\mu\| + \sup_{\lambda\in\mathcal{M}^+(\Omega_\varepsilon)} \langle \widetilde{g} - \mathbf{T}\mu, \lambda\rangle \right\} \\ &= \inf_{\substack{\mu\in\mathcal{M}^+(-\infty,K] \\ \mathbf{T}\mu\ge\widetilde{g}}} \|\mu\| = p_\varepsilon \end{aligned} \tag{2.73}$$

The reader easily verifies that for any $\lambda \in \mathcal{M}(\Omega_\varepsilon)$ the mapping $\mathcal{M}(-\infty, K] \ni \mu \mapsto K_\lambda(\mu) := K(\mu, \lambda)$ is closed in the sense of Section 5.4 and convex.

2.74 Lemma: *The* dual value function $v : C_0(-\infty, K] \mapsto (-\infty, \infty]$ *defined by*

$$v(f) := \inf_{\lambda\in\mathcal{M}(\Omega_\varepsilon)} K_\lambda^*(f)$$

is convex and we have $v(0) = -d_\varepsilon \ge v^{**}(0) = -p_\varepsilon$*. Here we denote by* K_λ^* *the conjugate of the mapping* K_λ*, see Equation* (5.22)*.*

Proof. By virtue of Lemma 5.25 and Equation (2.72) we find that

$$v^{**}(0) \leq v(0) = \inf_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} K_\lambda^*(0) = -\sup_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} \inf_{\mu \in \mathcal{M}(-\infty,K]} K(\mu,\lambda) = -d_\varepsilon.$$

The conjugate $v^* : \mathcal{M}(-\infty, K] \mapsto [-\infty, \infty]$ of the function v is given by

$$\begin{aligned} v^*(\mu) &= \sup_{f \in C_0(-\infty,K]} \{\langle f, \mu\rangle - v(f)\} \\ &= \sup_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} \sup_{f \in C_0(-\infty,K]} \{\langle f, \mu\rangle - K_\lambda^*(f)\} \\ &= \sup_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} K_\lambda^{**}(\mu) \\ &= \sup_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} K(\mu, \lambda). \end{aligned}$$

The last equality follows from Theorem 5.29 because the mapping K_λ is closed and convex. Hence, the biconjugate of the dual value function is given by

$$\begin{aligned} v^{**}(f) &= \sup_{\mu \in \mathcal{M}(-\infty,K]} \{\langle f, \mu\rangle - v^*(\mu)\} \\ &= \sup_{\mu \in \mathcal{M}(-\infty,K]} \inf_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} \{\langle f, \mu\rangle - K(\mu,\lambda)\} \end{aligned} \tag{2.75}$$

and owing to Equation (2.73), we obtain $v^{**}(0) = -p_\varepsilon$. Next, we show that the mapping v never assumes the value $-\infty$. Suppose there exists some $f \in C_0(-\infty, K]$ such that $v(f) = -\infty$. From Lemma 5.25 we know that $v^{**} \leq v$ and hence $v^{**}(f) = -\infty$ must hold true. Equation (2.75) now implies that

$$\sup_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} K(\mu, \lambda) = \infty$$

for any measure $\mu \in \mathcal{M}(-\infty, K]$ and therefore $p_\varepsilon = \infty$. This is impossible because the set of P_ε-admissible measures has already been shown to be non-empty.
In order to verify that v is convex, suppose that $\alpha \in (0,1)$ and $f_1, f_2 \in C_0(-\infty, K]$. From Equation (2.71) it is apparent that the Lagrange function K is concave in the second component and this yields

$$\begin{aligned} v(\alpha f_1 + (1-\alpha) f_2) &= \inf_{\lambda \in \mathcal{M}(\Omega_\varepsilon)} \sup_{\mu \in \mathcal{M}(-\infty,K]} \{\langle \alpha f_1 + (1-\alpha) f_2, \mu\rangle - K(\mu,\lambda)\} \\ &\leq \sup_{\mu \in \mathcal{M}(-\infty,K]} \{\langle \alpha f_1 + (1-\alpha) f_2, \mu\rangle - K(\mu, \alpha\lambda_1 + (1-\alpha)\lambda_2)\} \\ &\leq \alpha \sup_{\mu \in \mathcal{M}(-\infty,K]} \{\langle f_1, \mu\rangle - K(\mu, \lambda_1)\} \\ &\quad + (1-\alpha) \sup_{\mu \in \mathcal{M}(-\infty,K]} \{\langle f_2, \mu\rangle - K(\mu,\lambda_2)\} \end{aligned} \tag{2.76}$$

for any choice of $\lambda_0, \lambda_1 \in \mathcal{M}(\Omega_\varepsilon)$. Minimizing with respect to λ_0, λ_1 proves that v is indeed a convex function. □

Hence, strong duality holds if we can show that $v^{**}(0) = v(0)$ is true. By virtue of Lemma 2.74 and Theorem 5.29 we obtain $v^{**}(0) = \mathrm{cl}(\mathrm{co}(v))(0) = \mathrm{lsc}(v)(0)$. Using the representation of the lower semi-continuous hull from Lemma 5.25 yields

$$v^{**}(0) = \sup_{O \in \mathcal{U}(0)} \inf_{f \in O \setminus \{0\}} v(f)$$

where $\mathcal{U}(0)$ denotes the set containing all $\mathcal{T}_{\mathrm{uc}}$-open neighborhoods of 0. To put differently, in order to verify strong duality, it is sufficient to show that the mapping v is continuous at the origin with respect to the topology of uniform convergence. We will use the following adaptation of [AL, Theorem 5.42] to locally convex spaces.

2.77 Lemma: *Let V be a locally convex space, $f : V \to (-\infty, \infty]$ a convex function and $x_0 \in V$. If there exists an open neighborhood O of x_0 such that $\sup_{x \in O} f(x) < \infty$, then f is continuous at x_0.*

The set $O := \{\|f\|_\infty < 1\}$ is certainly a $\mathcal{T}_{\mathrm{uc}}$-open neighborhood of 0 and for any $f \in O$ we have

$$\begin{aligned} v(f) &= \inf_{\lambda \in \mathcal{M}^+(\Omega_\varepsilon)} \sup_{\mu \in \mathcal{M}^+(-\infty, K]} \{\langle f, \mu \rangle - \|\mu\| - \langle \widetilde{g}, \lambda \rangle + \langle \mathbf{T}\mu, \lambda \rangle\} \\ &\leq \sup_{\mu \in \mathcal{M}^+(-\infty, K]} \{\|\mu\| \|f\|_\infty - \|\mu\|\} = 0. \end{aligned}$$

Lemma 2.77 warrants that the mapping v is indeed continuous at 0 and therefore

$$p_\varepsilon = -v^{**}(0) = -v(0) = d_\varepsilon.$$

Next, we verify that the optimizers λ_ε and μ_ε satisfy the **complementary slackness** property. Taking the strong duality into account, we obtain

$$0 \leq \langle \mathbf{T}\mu_\varepsilon - \widetilde{g}, \lambda_\varepsilon \rangle = \langle \mathbf{T}\mu_\varepsilon, \lambda_\varepsilon \rangle - p_\varepsilon = \langle \mu_\varepsilon, \mathbf{T}^*\lambda_\varepsilon \rangle - d_\varepsilon = \langle \mu_\varepsilon, \mathbf{T}^*\lambda_\varepsilon - 1, \rangle \leq 0. \quad (2.78)$$

In other words, the equation $\mathbf{T}\mu_\varepsilon = \widetilde{g}$ holds λ_ε-a.e. on Ω_ε and $\mathbf{T}^*\lambda_\varepsilon = 1$ holds μ_ε-a.e. on $(-\infty, K]$. Moreover, the structure of the dual problem D_ε implies that we can always choose a D_ε-optimal element which assigns no mass to the zeros of the function $\widetilde{g}$, i.e. $\lambda_\varepsilon(\{(t, x) \in \Omega_\varepsilon \mid \widetilde{g}(x) = 0\}) = 0$. From now on we will only consider dual maximizers with this property. Let us summarize the findings from above:

2.79 Lemma: *For each $\varepsilon \in (0, \frac{T}{2})$ the linear programs P_ε, D_ε have solutions $\mu_\varepsilon, \lambda_\varepsilon$ and their optimal values $p_\varepsilon, d_\varepsilon$ coincide. The total mass of both optimizers is bounded by a constant $\rho \in \mathbb{R}_{++}$ which is independent of ε. Moreover, no mass of the measure λ_ε is located on the zero set of the function $\widetilde{g}$. The equation $\mathbf{T}\mu_\varepsilon = \widetilde{g}$ holds λ_ε-a.e. on Ω_ε and $\mathbf{T}^*\lambda_\varepsilon = 1$ holds μ_ε-a.e. on $(-\infty, K]$.*

We turn our attention to Program P_0 and the associated dual D_0 from Subsection 2.3.2. Theorem 2.63 will be proved in two steps. First, we will show that the primal optimizers $(\mu_\varepsilon)_{\varepsilon>0}$ cluster at some P_0-optimal measure μ_ε and that the family $(\lambda_\varepsilon)_{\varepsilon>0}$ contains a

D_0-admissible accumulation point λ_0. In the second step we will show that, under a certain additional requirement, the measure λ_0 is D_0-optimal. The other assertions of Theorem 2.63 will be verified along the way.

Step 1: Let p_0 and d_0 denote the optimal values of P_0 and D_0. The weak duality $0 \leq d_0 \leq p_0$ follows literally from the same calculation as in (2.68). Recall that for any $\varepsilon > 0$ the mass of the optimizers $\mu_\varepsilon \in \mathcal{M}^+(-\infty, K]$ and $\lambda_\varepsilon \in \mathcal{M}^+(\Omega)$ is bounded by some constant $\rho > 0$ independent of ε. General theory tells us that the vague topology is metrizable on the total variation unit balls in both spaces and Theorem 5.18 warrants that the latter constitute vaguely compact sets. Hence, we can find a sequence $\varepsilon_n \searrow 0$ and measures $\mu_0 \in \mathcal{M}^+(-\infty, K], \lambda_0 \in \mathcal{M}^+(\Omega)$ with $\|\mu_0\| \vee \|\lambda_0\| \leq \rho$ such that μ_{ε_n} converges vaguely to μ_0 and λ_{ε_n} converges vaguely to λ_0. For any $(t, x) \in \Omega$ the mapping $y \mapsto \kappa(t, x, y)$ is continuous on $(-\infty, K]$ and vanishes at infinity, see Equation (2.60). By vague convergence we conclude that

$$\mathbf{T}\mu_0(t,x) = \int_{-\infty}^{K} \kappa(t,x,y)\ \mathrm{d}\mu_0(y) = \lim_{n\to\infty} \int_{-\infty}^{K} \kappa(t,x,y)\ \mathrm{d}\mu_{\varepsilon_n}(y) \geq \widetilde{g}(x)$$

holds true for any $(t, x) \in \Omega$. This ensures that μ_0 is indeed P_0-admissible. Next, we verify that the measure λ_0 is D_0-admissible. Obviously, for any $\delta \in (0, T/4)$ we have $\emptyset \neq \Omega_{2\delta} \subset \Omega_\delta \subset \Omega$. By virtue of Urysohn's lemma, cf. [LA, p.40, Theorem 4.2], there exists a continuous function $\phi^\delta : \Omega \to [0, 1]$ such that $\phi^\delta(t, x) = 1$ for all $(t, x) \in \Omega_{2\delta}$ and $\phi^\delta(t, x) = 0$ for all $(t, x) \in \mathrm{cl}(\Omega \setminus \Omega_\delta)$. For any $y \in (-\infty, K]$ the continuous mapping $\Omega \ni (x, t) \mapsto \kappa(t, x, y)\phi^\delta(t, x)$ vanishes at infinity and by vague convergence of the sequence $\lambda_{\varepsilon_n} \to \lambda_0$ we obtain

$$\begin{aligned}
\int_\Omega \kappa(t,x,y)\mathrm{d}\lambda_0(t,x) &= \lim_{\delta\searrow 0} \int_\Omega \kappa(t,x,y)\mathbb{1}_{\Omega_{2\delta}}(t,x)\ \mathrm{d}\lambda_0(t,x)\\
&\leq \lim_{\delta\searrow 0} \int_\Omega \kappa(t,x,y)\phi^\delta(t,x)\ \mathrm{d}\lambda_0(t,x)\\
&= \lim_{\delta\searrow 0} \lim_{n\to\infty} \int_\Omega \kappa(t,x,y)\phi^\delta(t,x)\ \mathrm{d}\lambda_{\varepsilon_n}(t,x)\\
&\leq \limsup_{\delta\searrow 0} \limsup_{n\to\infty} \int_\Omega \kappa(t,x,y)\ \mathrm{d}\lambda_{\varepsilon_n}(t,x) \leq 1.
\end{aligned}$$

In other words, the measure λ_0 is dual admissible. Next, let us establish the strong duality $p_0 = d_0$ by putting together several of the previous results. The vague convergence of the measures μ_{ε_n} to μ_0 implies that $\|\mu_0\| \leq \liminf_{n\to\infty} \|\mu_{\varepsilon_n}\|$ is true. Recalling that strong duality holds in the P_ε-D_ε-setting now yields

$$d_0 \leq p_0 \leq \|\mu_0\| \leq \liminf_{n\to\infty} \|\mu_{\varepsilon_n}\| = \liminf_{n\to\infty} p_{\varepsilon_n} = \liminf_{n\to\infty} d_{\varepsilon_n} \leq d_0. \tag{2.80}$$

The last inequality follows from the fact that all D_ε-admissible elements are certainly D_0-admissible. Along the way we proved that the P_0-optimal value is attained by the measure μ_0. Additionally, we observe that any P_0-admissible element assigns mass to every open subset of $(-\infty, K)$. Indeed, assuming that the latter is false allows us to pick some P_0-admissible measure μ and a bounded, open interval $I := (c-\nu, c+\nu) \subset (-\infty, K)$ such that $\mu(I) = 0$. Obviously, we have $0 < \delta := \inf_{x \in I} \widetilde{g}(x)$ and this yields

$$\delta < \widetilde{g}(c) \leq \mathbf{T}\mu(t,c) = \int_{-\infty}^{K} \mathbb{1}_{\{|y-c| \geq \nu\}} \, \kappa(t,c,y) \, \mathrm{d}\mu(y) \tag{2.81}$$

for all $t \in (0,T)$. In consideration of (2.60), we find that the right-hand side of Equation (2.81) converges to 0 as $t \nearrow T$. This contradiction proves the claim.

Step 2: We will show that the D_0-optimal value is attained by λ_0 if some additional requirement is met. We already know that the measure λ_0 is D_0-admissible and that the sequence λ_{ε_n} converges to λ_0 with respect to the vague topology on $\mathcal{M}(\Omega)$. Due to the lack of compactness, we cannot directly conclude that λ_{ε_n} converges weakly to λ_0. Please, observe that the functional $\mathcal{M}(\Omega) \ni \lambda \mapsto \langle \widetilde{g}, \lambda \rangle$ is weakly but not vaguely continuous.

First, we prove that the sequence λ_{ε_n} converges weakly in $\mathcal{M}(\operatorname{cl}\Omega)$ where $\operatorname{cl}\Omega = [0,T] \times (-\infty, K]$. It is sufficient to show that the family $\{\lambda_{\varepsilon_n} \mid n \in \mathbb{N}\}$ is tight. For each $\varepsilon > 0$ we define by $K_\varepsilon := [0,T] \times [-1/\varepsilon, K]$ a compact subset of $\operatorname{cl}\Omega$. For every $n \in \mathbb{N}$ the mass of λ_{ε_n} is concentrated on $\Omega_{\varepsilon_n} \subset K_{\varepsilon_n}$. Let us assume that the family of measures is not tight. Hence, there exists a constant $\delta > 0$ such that for any $n \in \mathbb{N}$ we can choose some integer $M(n) \geq n$ such that

$$\lambda_{\varepsilon_{M(n)}}(\Omega \setminus K_{\varepsilon_n}) > \delta.$$

Now pick a sufficiently small constant $C \in (-\infty, K)$ such that

$$\int_{-\infty}^{C} \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right) \, \mathrm{d}y \leq \frac{\delta}{2}.$$

Due to the fact that all measures λ_{ε_n} are D_0-admissible, we have

$$\int_{\Omega} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}\lambda_{\varepsilon_{M(n)}}(t,x) \leq \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)$$

for any $y \in (-\infty, K]$. Integrating the latter inequality over the set $(-\infty, C)$ yields

$$\int_{\Omega} \int_{-\infty}^{C} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}y \, \mathrm{d}\lambda_{\varepsilon_{M(n)}}(t,x) \leq \frac{\delta}{2}.$$

Due to the positivity of measure and integrand, we conclude that

$$\begin{aligned}
\frac{\delta}{2} &\geq \int_{\Omega \setminus K_{\varepsilon_n}} \int_{-\infty}^{C} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}y \, \mathrm{d}\lambda_{\varepsilon_{M(n)}}(t,x) \\
&\geq \lambda_{\varepsilon_{M(n)}}\left(\Omega \setminus K_{\varepsilon_n}\right) \inf_{(t,x) \in \Omega \setminus K_{\varepsilon_n}} \int_{-\infty}^{C} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y\right) \, \mathrm{d}y \\
&\geq \delta \inf_{x < -1/\varepsilon_n} \inf_{t \in [0,T]} \int_{-\infty}^{C} \mathrm{N}\left(x + \tilde{r}t, \sigma^2 t, y\right) \, \mathrm{d}y
\end{aligned}$$

holds true for all $n \in \mathbb{N}$. Taking the limit $n \to \infty$ yields

$$\frac{\delta}{2} \geq \delta \lim_{n\to\infty} \inf_{x<-1/\varepsilon_n} \inf_{t\in[0,T]} \int_{-\infty}^{C} \mathrm{N}\left(x + \tilde{r}t, \sigma^2 t, y\right) \,\mathrm{d}y = \delta$$

as $\varepsilon_n \to 0$. This is impossible and consequently the family $\{\lambda_{\varepsilon_n} \mid n \in \mathbb{N}\}$ must be tight. Hence, the sequence λ_{ε_n} converges weakly in $\mathcal{M}(\mathrm{cl}\,\Omega)$ to some measure $\overline{\lambda}_0$ with $\overline{\lambda}_0|_\Omega = \lambda_0$. It is sufficient to show that $\overline{\lambda}_0$ assigns no mass to the borders $M_1 := \{0\} \times (-\infty, K)$ and $M_2 := \{T\} \times (-\infty, K)$ in order to assure that the measure λ_0 is D_0-optimal. Indeed, in this case we find that

$$\int_\Omega \tilde{g}(x) \,\mathrm{d}\lambda_0(t,x) = \int_{\mathrm{cl}\,\Omega} \tilde{g}(x) \,\mathrm{d}\overline{\lambda}_0(t,x) = \lim_{n\to\infty} \int_\Omega \tilde{g}(x) \,\mathrm{d}\lambda_{\varepsilon_n}(t,x) = \lim_{n\to\infty} d_{\varepsilon_n} = d_0$$

is true. The second equality follows from the weak convergence of the sequence λ_{ε_n} in the space $\mathcal{M}(\mathrm{cl}\,\Omega)$ and the boundedness of the continuous function $\tilde{g}$. The last equality has already been established in Equation (2.80).

First, assume that $\overline{\lambda}_0$ assigns mass to the set M_1. In this case we can choose a real number $\alpha < K$ such that $\overline{\lambda}_0\left(\{0\} \times [\alpha, K)\right) > 0$. Due to weak convergence in the space $\mathcal{M}\left([0, T/2] \times [\alpha, K]\right)$, we find that

$$\begin{aligned}
\int_{\{0\}\times[\alpha,K)} \kappa(t,x,y) \,\mathrm{d}\overline{\lambda}_0(t,x) &\leq \int_{[0,T/2]\times[\alpha,K]} \kappa(t,x,y) \,\mathrm{d}\overline{\lambda}_0(t,x) \\
&= \lim_{n\to\infty} \int_{[0,T/2]\times[\alpha,K]} \kappa(t,x,y) \,\mathrm{d}\lambda_{\varepsilon_n}(t,x) \leq 1
\end{aligned}$$

holds true for any $y \in (-\infty, K]$. Fatou's lemma and Lemma 5.2 now yield the following contradiction

$$\begin{aligned}
1 &\geq \liminf_{y\to-\infty} \int_{\{0\}\times[\alpha,K)} \frac{\mathrm{N}\left(x + \tilde{r}T, \sigma^2 T, y\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)} \,\mathrm{d}\overline{\lambda}_0(t,x) \\
&\geq \int_{\{0\}\times[\alpha,K)} \liminf_{y\to-\infty} \exp\left(y \frac{x - x_0}{\sigma^2 T}\right) \exp\left(\frac{x_0^2 - x^2 + 2\tilde{r}T(x_0 - x)}{2\sigma^2 T}\right) \,\mathrm{d}\overline{\lambda}_0(t,x) = \infty
\end{aligned}$$

as $K < x_0$ by choice. Our assumption was wrong and therefore $\overline{\lambda}_0(M_1) = 0$. Next, we turn our attention to the set M_2. For any $(t,x) \in [0,T] \times \mathbb{R}$ we define by

$$V(t,x) := \liminf_{\substack{(t',x')\to(t,x) \\ (t',x')\in(0,T)\times\mathbb{R}}} \mathbf{T}\mu_0(t',x') \tag{2.82}$$

the lower semi-continuous extension of the function $\mathbf{T}\mu_0$ to the set $[0,T] \times \mathbb{R}$. We will show that imposing the *additional assumption*

$$V(T,x) > \tilde{g}(x) \quad \forall x \in (-\infty, K) \tag{2.83}$$

warrants that the measure $\overline{\lambda}_0$ assigns no mass to the set M_2. Indeed, by virtue of Lemma 5.9 we find that V attains its minimal value on any compact subset of $[0,T] \times \mathbb{R}$.

Moreover, assumption (2.83) ensures that the minimal value of the function $V - \widetilde{g}$ is strictly positive on any set of the type $\{T\} \times [a, b] \subset M_2$ where $a < b < K$. By lower semi-continuity we can choose some $n_0 \in \mathbb{N}$ and $\delta > 0$ such that

$$V(t, x) - \widetilde{g}(x) \geq \delta \tag{2.84}$$

for any $(t, x) \in [T - 1/n_0, T] \times [a, b]$. Now assume that the measure $\overline{\lambda}_0$ assigns mass to M_2. We can choose some strip $\{T\} \times (a, b) \subset M_2$ and a constant $\rho > 0$ such that $\overline{\lambda}_0(Q_m) \geq 2\rho$ holds for any $m \in \mathbb{N}$ where $Q_m := (T - 1/m, T] \times (a, b)$. The measures λ_{ε_n} converge weakly in $\mathcal{M}(\operatorname{cl}\Omega)$ to $\overline{\lambda}_0$ and owing to [KL, Theorem 13.16], we can pass to a subsequence (again denoted by ε_n) such that $\lambda_{\varepsilon_n}(\operatorname{int} Q_n) \geq \rho$ for all $n \in \mathbb{N}$. The strong duality in the D_ε-P_ε-setting allows us to conclude that

$$\begin{aligned}
\langle \mathbf{T}\mu_0 - \widetilde{g}, \lambda_{\varepsilon_n} \rangle &= \langle \mu_0, \mathbf{T}^*\lambda_{\varepsilon_n} \rangle - p_{\varepsilon_n} \\
&= \langle \mu_0, \mathbf{T}^*\lambda_{\varepsilon_n} - 1 \rangle + \|\mu_0\| - \|\mu_{\varepsilon_n}\| \\
&\leq \|\mu_0\| - \|\mu_{\varepsilon_n}\|.
\end{aligned}$$

Moreover, Equation (2.84) implies that

$$\begin{aligned}
\langle \mathbf{T}\mu_0 - \widetilde{g}, \lambda_{\varepsilon_n} \rangle &\geq \int_{\operatorname{int} Q_n} V(t, x) - \widetilde{g}(x) \, \mathrm{d}\lambda_{\varepsilon_n}(t, x) \\
&\geq \delta \; \lambda_{\varepsilon_n}(\operatorname{int} Q_n) \\
&\geq \delta \; \rho > 0
\end{aligned}$$

holds true for any integer $n \geq n_0$. And yet, we already know from Equation (2.80) that $\|\mu_0\| - \|\mu_{\varepsilon_n}\| \to 0$ as $n \to \infty$. This yields a contradiction which finally shows that $\overline{\lambda}_0(M_2) = 0$. Last but not least, we observe that literally the same calculation as in (2.78) yields the complementary slackness property for μ_0 and λ_0 in the case of primal and dual attainment. This means that the equation

$$\int_{-\infty}^{K} \kappa(t, x, y) \, \mathrm{d}\mu_0(y) = \widetilde{g}(x) \tag{2.85}$$

holds λ_0-a.e. on Ω and

$$\int_{\Omega} \kappa(t, x, y) \, \mathrm{d}\lambda_0(t, x) = 1 \tag{2.86}$$

holds μ_0-a.e. on $(-\infty, K]$. Let us summarize our results from above.

2.87 Theorem:

1. *For any $\varepsilon \in (0, \frac{T}{2})$ the linear programs P_ε and D_ε have solutions μ_ε and λ_ε. The optimal values p_ε and d_ε of the latter programs coincide. The total mass of the optimizers is bounded by some constant $\rho \in \mathbb{R}_{++}$ which does not depend on ε. Moreover, the measure λ_ε assigns no mass to the zero set of the function $\widetilde{g}$. The equation $\mathbf{T}\mu_\varepsilon = \widetilde{g}$ holds λ_ε-a.e. on Ω_ε and $\mathbf{T}^*\lambda_\varepsilon = 1$ holds μ_ε-a.e. on $(-\infty, K]$.*

2. *There exists a sequence $\varepsilon_n \searrow 0$ such that μ_{ε_n} converges vaguely in $\mathcal{M}(-\infty, K]$ to some P_0-admissible measure μ_0 and λ_{ε_n} converges vaguely in $\mathcal{M}(\Omega)$ to some D_0-admissible measure λ_0. The optimal value of P_0 is obtained by μ_0 and coincides with the optimal value of D_0. The measure μ_0 assigns mass to every open subset of $(-\infty, K)$ and $\|\mu_0\| \vee \|\lambda_0\| \leq \rho$.*

3. *Let V be defined as in Equation (2.82). If $V(T,x) > \tilde{g}(x)$ for any $x \in (-\infty, K)$, the optimal value of the program D_0 is obtained by λ_0. In this case the complementary slackness equations (2.85) and (2.86) hold.*

Theorem 2.63 is nothing but a slight reformulation of the latter result.

2.4 Verification - proof of Theorem 2.52

We use the notation from the preceding sections. In particular, see Subsection 2.3.2 for the definition of the operator $\mathbf{T}$ and the formulation of the optimization problem P_0. Let μ^* be a cheapest dominating European option in the sense of Theorem 2.51. In light of Equation (2.58), we have

$$\begin{aligned} v_{\mathrm{eu},\mu^*}(\vartheta, x) &= e^{-r\vartheta} \int_{-\infty}^{K} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \mathrm{d}\mu^*(y) \\ &= e^{-\frac{2r}{\sigma^2}x} \int_{-\infty}^{K} \frac{\mathrm{N}\left(x + \tilde{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y\right)} e^{(2r/\sigma^2)x_0 - rT} \mathrm{d}\mu^*(y) \\ &= e^{-\frac{2r}{\sigma^2}x}\, \mathbf{T}\mu_0(T - \vartheta, x) \end{aligned} \tag{2.88}$$

for any $(\vartheta, x) \in (0, T) \times \mathbb{R}$. Here we denote by $\mu_0 = e^{(2r/\sigma^2)x_0 - rT}\mu^*$ the corresponding P_0-optimal measure from Theorem 2.63.

Step 1: Analyticity of the European value function
First, we show that the first assumption of Theorem 2.52 ensures the analyticity of the function v_{eu,μ^*} on the open $\mathbb{C}^2$-domain

$$E := \left\{ \vartheta \in \mathbb{C} \;\middle|\; \sqrt{(\operatorname{Re}\vartheta - (T + 2\delta)/2)^2 + (\operatorname{Im}\vartheta)^2} < (T + 2\delta)/2 \right\} \times \mathbb{C}.$$

Clearly, it is enough to verify that the function

$$e^{r\vartheta} v_{\mathrm{eu},\mu^*}(\vartheta, x) = \int_{-\infty}^{\infty} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{\mathrm{N}\left(x_1 + \hat{r}(T + 2\delta), \sigma^2(T + 2\delta), y\right)} \mathrm{d}\mu^{**}(y) \tag{2.89}$$

is analytic on E where

$$\frac{\mathrm{d}\mu^{**}}{\mathrm{d}\mu^*}(y) := \frac{\mathrm{N}\left(x_1 + \hat{r}(T + 2\delta), \sigma^2(T + 2\delta), y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)}.$$

In light of Assumption 1, we have $\|\mu^{**}\| = v_{\mathrm{eu},\mu^*}(T+2\delta, x_1)e^{r(T+2\delta)} < \infty$. Due to Hartogs' theorem, it is enough to show that the function from Equation (2.89) is partially analytic, cf. [KR, Paragraph 2.4]. Lemma 5.2 implies that

$$\begin{aligned}\left|\frac{\mathrm{N}(x+\hat{r}\vartheta, \sigma^2\vartheta, y)}{\mathrm{N}(x_1+\hat{r}(T+2\delta), \sigma^2(T+2\delta), y)}\right| &= |h_1(\vartheta,x)| \left|\exp\Big(-\frac{(y-A)^2}{2B}\Big)\right| \\ &= |h_2(\vartheta,x)| \exp\Big(-\frac{\operatorname{Re}B}{2|B|^2}\Big(y-\operatorname{Re}A-\frac{\operatorname{Im}A\operatorname{Im}B}{\operatorname{Re}B}\Big)^2\Big)\end{aligned}$$

for any $(\vartheta, x) \in E$ and $y \in \mathbb{R}$. Here we denote by h_1, h_2 certain functions which are continuous on E. The quantities A and B are defined as in Lemma 5.2. For any $(\vartheta, x) \in E$ we have

$$\operatorname{Re}B(\vartheta) = \operatorname{Re}\frac{\sigma^2\vartheta(T+2\delta)}{(T+2\delta-\vartheta)} = \frac{\sigma^2(T+2\delta)}{|T+2\delta-\vartheta|^2}\Big((T+2\delta)\operatorname{Re}\vartheta - |\vartheta|^2\Big) > 0$$

and therefore the integrand occurring in (2.89) satisfies the inequality

$$\sup_{y\in\mathbb{R}}\left|\frac{\mathrm{N}(x+\hat{r}\vartheta, \sigma^2\vartheta, y)}{\mathrm{N}(x_1+\hat{r}(T+2\delta), \sigma^2(T+2\delta), y)}\right| \le |h_2(\vartheta,x)|\,.$$

The quantity on the right-hand side is certainly bounded on every compact set contained in E. Hence, we can use a standard argument which involves the theorems of Morera and Fubini in order to prove partial analyticity. For a detailed exposition of the technique, we refer the reader to the proof of Lemma 5.3. In virtue of Hartogs' theorem we conclude that the mapping v_{eu,μ^*} is indeed analytic on E.

Step 2: Proof of Assertion (a)
We want to show that the curve $\vartheta \mapsto \widehat{x}(\vartheta)$ is analytic on an open complex domain containing the interval $(0, T]$. For this purpose we will use a version of the implicit function theorem from multivariate complex analysis which can be found in Section 5.2. By virtue of Step 1 and the assumptions imposed on the American payoff g, we observe that the function

$$\Psi(\vartheta, x) := v_{\mathrm{eu},\mu^*}(\vartheta, x) - g(x) \tag{2.90}$$

is analytic on the open $\mathbb{C}^2$-domain $D' \times D$ where

$$D' := \Big\{\vartheta \in \mathbb{C} \;\Big|\; \sqrt{(\operatorname{Re}\vartheta - (T+2\delta)/2)^2 + (\operatorname{Im}\vartheta)^2} < (T+2\delta)/2\Big\}$$

and D denotes the domain of analyticity of g. Clearly, the set D' is simply connected and $(0, T+2\delta) \times (-\infty, K)$ is a subset of $D' \times D$. The continuity of Ψ in combination with the uniqueness of the minima warrants that the curve $\widehat{x}$ is continuous on the interval $(0, T+\delta/2)$. Indeed, assume that $\vartheta_0 \in (0, T+\delta/2)$ is a point of discontinuity. Then we can choose a sequence $\vartheta_n \to \vartheta_0$ and some $x_\infty \le K, \varepsilon > 0$ such that $\widehat{x}(\vartheta_n) \to x_\infty$ as $n \to \infty$ and $|x_\infty - \widehat{x}(\vartheta_0)| > \varepsilon$. Clearly, there exists a constant $\gamma > 0$ such that

$\Psi(\vartheta_0, \widehat{x}(\vartheta_0)) + \gamma < \Psi(\vartheta_0, x_\infty)$. Consequently, we can choose two disjoint balls $B_r(\vartheta_0, x_\infty)$ and $B_r(\vartheta_0, \widehat{x}(\vartheta_0))$ of radius $r \in (0, \varepsilon/2)$ such that

$$\Psi(\vartheta, x) + \frac{\gamma}{2} < \Psi(\tilde{\vartheta}, \tilde{x})$$

holds true for any $(\vartheta, x) \in B_r(\vartheta_0, \widehat{x}(\vartheta_0))$ and $(\tilde{\vartheta}, \tilde{x}) \in B_r(\vartheta_0, x_\infty)$. This yields a contradiction as $(\vartheta_n, \widehat{x}(\vartheta_n))$ is contained in $B_r(\vartheta_0, x_\infty)$ for any sufficiently large integer n. Hence, the curve $\widehat{x}$ must be continuous.

Moreover, for any $\vartheta \in (0, T + \delta)$ we have $\partial_x \Psi(\vartheta, \widehat{x}(\vartheta)) = 0$ and $\partial_{xx} \Psi(\vartheta, \widehat{x}(\vartheta)) \geq 0$ due to the necessary first and second order optimality conditions. Applying Kolmogorov's backward equation, we obtain

$$\begin{aligned}
\partial_{xx}\Psi &= \partial_{xx} v_{\mathrm{eu},\mu^*} - g'' \\
&= \frac{2}{\sigma^2}\partial_\vartheta v_{\mathrm{eu},\mu^*} + \left(1 - \frac{2r}{\sigma^2}\right)\partial_x v_{\mathrm{eu},\mu^*} + \frac{2r}{\sigma^2} v_{\mathrm{eu},\mu^*} - g'' \\
&= \frac{2}{\sigma^2}\partial_\vartheta v_{\mathrm{eu},\mu^*} + \left(1 - \frac{2r}{\sigma^2}\right)\partial_x \Psi + \frac{2r}{\sigma^2}\left(v_{\mathrm{eu},\mu^*} - g\right) - c \\
&= H + \left(1 - \frac{2r}{\sigma^2}\right)\partial_x \Psi
\end{aligned} \tag{2.91}$$

on $(0, T + 2\delta) \times \mathbb{R}$ where the quantities c and H are defined as in (2.44) and (2.53), respectively. Assumption 3 of the theorem warrants that H and therefore $\partial_{xx}\Psi$ is strictly positive on the set $\Gamma := \{(\vartheta, \widehat{x}(\vartheta)) \mid \vartheta \in (0, T]\}$ and therefore Theorem 5.7 is applicable to the function $\partial_x \Psi$ at any point of Γ. We obtain that for any $(\tilde{\vartheta}, \tilde{x}) \in \Gamma$ there exist open neighborhoods $\tilde{\vartheta} \in U_{\tilde{\vartheta}}, \tilde{x} \in U_{\tilde{x}}$ and an analytic curve $\chi_{\tilde{\vartheta}} : U_{\tilde{\vartheta}} \to U_{\tilde{x}}$ such that $\chi_{\tilde{\vartheta}}(\vartheta) = \widehat{x}(\vartheta)$ holds true for any $\vartheta \in U_{\tilde{\vartheta}} \cap (0, T]$. The identity theorem implies that curves $\chi_{\tilde{\vartheta}_1}, \chi_{\tilde{\vartheta}_2}$ with intersecting neighborhoods $U_{\tilde{\vartheta}_1}, U_{\tilde{\vartheta}_2}$ coincide on $U_{\tilde{\vartheta}_1} \cup U_{\tilde{\vartheta}_2}$. From above we already know that the mapping $\vartheta \mapsto \widehat{x}(\vartheta)$ is continuous and consequently there exists an analytic function χ such that $\chi|_{U_{\tilde{\vartheta}}} = \chi_{\tilde{\vartheta}}$ for any neighborhood $U_{\tilde{\vartheta}}$. In particular, we have $\chi(\vartheta) = \widehat{x}(\vartheta)$ for every $\vartheta \in (0, T]$. This proves that $\widehat{x}$ is indeed analytic on some complex domain containing the interval $(0, T]$.

Step 3: Proof of Assertion (b)

Now we will verify that $v_{\mathrm{eu},\mu^*}(\vartheta, \widehat{x}(\vartheta)) = g(\widehat{x}(\vartheta))$ holds true for all $\vartheta \in [0, T]$. Due to the fact that the measure μ^* assigns no mass to the set (K, ∞), we find that $v_{\mathrm{eu},\mu^*}(0, x) = 0$ for any $x > K$. Lower semi-continuity even implies that $v_{\mathrm{eu},\mu^*}(0, K) = 0$ and therefore $v_{\mathrm{eu},\mu^*}(0, \widehat{x}(0)) = v_{\mathrm{eu},\mu^*}(0, K) = 0 = g(K) = g(\widehat{x}(0))$. In light of Equation (2.88), we have

$$e^{\frac{2r}{\sigma^2}x}\left(v_{\mathrm{eu},\mu^*}(T - t, x) - g(x)\right) = \mathbf{T}\mu_0(t, x) - \widetilde{g}(x)$$

for any $(t, x) \in [0, T) \times \mathbb{R}$ where $\widetilde{g}$ is defined as in (2.57). Assumption 2 implies that $v_{\mathrm{eu},\mu^*}(0, x) - g(x) > 0$ for every $x < K$ and consequently Theorem 2.87 warrants strong duality, primal and dual attainment as well as complementary slackness. By virtue of the complementary slackness equation (2.85), we find that the dual maximizer λ_0 assigns no mass to the complement of the set $\{(t, \widehat{x}(T-t)) \mid 0 < t < T\}$. We claim that there exists a sequence $\vartheta_n \nearrow T$ with $\vartheta_n \in (0,T)$ such that

$$v_{\mathrm{eu},\mu^*}(\vartheta_n, \widehat{x}(\vartheta_n)) = g(\widehat{x}(\vartheta_n)) \tag{2.92}$$

for any $n \in \mathbb{N}$. Assume the latter statement is false. Then we can pick some $\varepsilon \in (0,T)$ such that $v_{\mathrm{eu},\mu^*}(\vartheta, \widehat{x}(\vartheta)) > g(\widehat{x}(\vartheta))$ for all $\vartheta \in (T-\varepsilon, T)$. Equation (2.85) tells us that the measure λ_0 is concentrated on the set $\Gamma_\varepsilon := \{(t, x(T-t)) \mid \varepsilon < t < T\}$. From Theorem 2.87 we already know that the primal minimizer μ_0 assigns mass to every open subset of $(-\infty, K)$. In light of Equation (2.86), we can find a sequence $y_n \searrow -\infty$ with $\max_{n\in\mathbb{N}} y_n < \min_{\vartheta\in[0,T]} \widehat{x}(\vartheta) + \tilde{r}(T-\varepsilon) =: \omega_0$ such that

$$\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y_n\right) = \int_{\Gamma_\varepsilon} \mathrm{N}\left(x + \tilde{r}(T-t), \sigma^2(T-t), y_n\right) \, \mathrm{d}\lambda_0(t,x)$$

holds true for all $n \in \mathbb{N}$. Due to the fact that $\tilde{r} < 0$, we have

$$\begin{aligned} \mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y_n\right) &\leq \int_{\Gamma_\varepsilon} \mathrm{N}\left(\omega_0, \sigma^2(T-t), y_n\right) \, \mathrm{d}\lambda_0(t,x) \\ &\leq \mathrm{N}\left(\omega_0, \sigma^2(T-\varepsilon), y_n\right) \, \lambda_0\left(\Gamma_\varepsilon\right) \end{aligned}$$

for any $n \in \mathbb{N}$. This yields the contradiction

$$1 \leq \lambda_0\left(\Gamma\right) \lim_{n\to\infty} \frac{\mathrm{N}\left(\omega_0, \sigma^2(T-\varepsilon), y_n\right)}{\mathrm{N}\left(x_0 + \tilde{r}T, \sigma^2 T, y_n\right)} = 0$$

and consequently a sequence with the desired property (2.92) must exist. In light of Step 1 and Step 2 from above, we find that the mapping $\vartheta \mapsto v_{\mathrm{eu},\mu^*}(\vartheta, \widehat{x}(\vartheta)) - g(\widehat{x}(\vartheta))$ is analytic on some open complex domain containing the interval $(0,T]$. Equation (2.92) and the identity theorem finally yield that $v_{\mathrm{eu},\mu^*}(\vartheta, \widehat{x}(\vartheta)) = g(\widehat{x}(\vartheta))$ for any $\vartheta \in (0,T]$.

Step 4: Proof of Assertion (c)
Now we will verify that μ^* is the unique measure representing our American payoff on the set $\widetilde{C}_{(T,x_0)}$ as defined in (2.54). Moreover, we will show that $\widetilde{C}_{(T,x_0)}$ constitutes a connected subset of the continuation region's connected component $C_{(T,x_0)}$, cf. Subsection 2.1.1. Not surprisingly, for any $T_0 \in [0,T]$ the process $V_t^{T_0} := e^{-rt} v_{\mathrm{eu},\mu^*}(T_0 - t, X_t)$ is a martingale on the time segment $[0, T_0)$. Indeed, for $0 \leq u < t+u < T_0$ the Markov property of the process X yields

$$\begin{aligned} \mathbb{E}_x\left[V_{t+u}^{T_0} \middle| \mathcal{F}_u\right] &= e^{-r(t+u)} \mathbb{E}_{X_u}\left[v_{\mathrm{eu},\mu^*}(T_0 - t - u, X_t)\right] \\ &= e^{-rT_0} \int_{-\infty}^{K} \frac{\mathbb{E}_{X_u}\left[\mathrm{N}\left(X_t + \hat{r}(T_0 - t - u), \sigma^2(T_0 - t - u), y\right)\right]}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \, \mathrm{d}\mu^*(y) \\ &= e^{-rT_0} \int_{-\infty}^{K} \frac{\mathrm{N}\left(X_u + \hat{r}(T_0 - u), \sigma^2(T_0 - u), y\right)}{\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)} \, \mathrm{d}\mu^*(y) \\ &= e^{-ru} v_{\mathrm{eu},\mu^*}(T_0 - u, X_u) = V_u^{T_0}. \end{aligned} \tag{2.93}$$

The third equality follows from the convolution property of the normal distribution. The martingale condition cannot be assured at the end time of the market. Nevertheless, Fatou's lemma yields the super-martingale property. Indeed, for any $u \in [0, T_0]$ we have

$$\begin{aligned}\mathbb{E}_x\left[V_{T_0}^{T_0}\middle|\mathcal{F}_u\right] &= \mathbb{E}_x\left[e^{-rT_0}\liminf_{t\nearrow T_0} v_{\mathrm{eu},\mu^*}(T_0-t, X_t)\middle|\mathcal{F}_u\right]\\ &\le \liminf_{t\nearrow T_0}\mathbb{E}_x\left[e^{-rt}v_{\mathrm{eu},\mu^*}(T_0-t, X_t)\middle|\mathcal{F}_u\right] = V_u^{T_0}.\end{aligned}$$

Due to superreplication, we have $e^{-rt}g(X_t) \le V_t^T$ for any $t \in [0, T]$ and consequently the optional sampling theorem yields

$$\mathbb{E}_x\left[e^{-r\tau}g(X_\tau)\middle|\mathcal{F}_t\right] \le \mathbb{E}_x\left[V_\tau^T\middle|\mathcal{F}_t\right] \le V_t$$

for any $[t, T]$-valued stopping time τ. Maximizing the left-hand side over all such stopping times shows that $v_{\mathrm{am},g}(\vartheta, x) \le v_{\mathrm{eu},\mu^*}(\vartheta, x)$ for any $(\vartheta, x) \in [0, T] \times \mathbb{R}$.

Next, we will prove that the value functions $v_{\mathrm{am},g}$ and v_{eu,μ^*} coincide on the set $\widetilde{C}_{(T,x_0)}$. For this purpose let τ_ϑ be defined as in (2.55). Assumption 4 warrants that the measure μ^* has no point mass at K, i.e. $\mu^*(\{K\}) = 0$. Indeed, assuming $\mu^*(\{K\}) > 0$ would imply that

$$\liminf_{\vartheta\to 0} v_{\mathrm{eu},\mu^*}(\vartheta, K) \ge \gamma \liminf_{\vartheta\to 0} e^{-r\vartheta}\mathrm{N}\left(\hat{r}\vartheta, \sigma^2\vartheta, 0\right) = \infty$$

where γ denotes some positive constant. Furthermore, owing the geometric properties of the curve $\widehat{x}$, we have

$$\mathbb{E}_x\left[\mathrm{N}\left(X_{\tau_\vartheta} + \hat{r}(\vartheta-\tau_\vartheta), \sigma^2(\vartheta-\tau_\vartheta), y\right)\mathbb{1}_{\{\tau_\vartheta=\vartheta\}}\right] \le \mathbb{E}_x\left[\delta_{\{y\}}(X_\vartheta)\,\mathbb{1}_{\{X_\vartheta\ge K\}}\right] = 0$$

for any $y < K$. Hence, for any $(\vartheta, x) \in \widetilde{C}_{(T,x_0)}$ we obtain by monotone convergence

$$\begin{aligned}v_{\mathrm{am},g}(\vartheta, x) &\ge \mathbb{E}_x\left[e^{-r\tau_\vartheta}g(X_{\tau_\vartheta})\,\mathbb{1}_{\{\tau_\vartheta<\vartheta\}}\right]\\ &= \mathbb{E}_x\left[e^{-r\tau_\vartheta}v_{\mathrm{eu},\mu^*}(\vartheta-\tau_\vartheta, X_{\tau_\vartheta})\,\mathbb{1}_{\{\tau_\vartheta<\vartheta\}}\right]\\ &= \lim_{x'\nearrow K}\mathbb{E}_x\left[e^{-r\vartheta}\int_{-\infty}^{x'}\frac{\mathrm{N}\left(X_{\tau_\vartheta}+\hat{r}(\vartheta-\tau_\vartheta), \sigma^2(\vartheta-\tau_\vartheta), y\right)}{\mathrm{N}\left(x_0+\hat{r}T, \sigma^2T, y\right)}\mathrm{d}\mu^*(y)\,\mathbb{1}_{\{\tau_\vartheta<\vartheta\}}\right]\\ &= \lim_{x'\nearrow K} e^{-r\vartheta}\int_{-\infty}^{x'}\frac{\mathbb{E}_x\left[\mathrm{N}\left(X_{\tau_\vartheta}+\hat{r}(\vartheta-\tau_\vartheta), \sigma^2(\vartheta-\tau_\vartheta), y\right)\right]}{\mathrm{N}\left(x_0+\hat{r}T, \sigma^2T, y\right)}\mathrm{d}\mu^*(y)\\ &= v_{\mathrm{eu},\mu^*}(\vartheta, x).\end{aligned}$$

Summing up, we have shown that $v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},\mu^*}(\vartheta, x) > g(x)$ holds true for any $(\vartheta, x) \in \widetilde{C}_{(T,x_0)}$. Besides, the latter directly implies that $\widetilde{C}_{(T,x_0)}$ is a connected subset of $C_{(T,x_0)}$.

Finally, we prove that the representing measure μ^* is unique. Assume that we can find another measure ν such that $v_{\mathrm{eu},\mu^*}(\vartheta, x) = v_{\mathrm{am},g}(\vartheta, x) = v_{\mathrm{eu},\nu}(\vartheta, x)$ holds true for any

$(\vartheta, x) \in \widetilde{C}_{(T,x_0)}$. From above we know that the value functions $v_{\mathrm{eu},\mu^*}, v_{\mathrm{eu},\nu}$ are analytic on a $\mathbb{C}^2$-domain containing the set $(0,T) \times \mathbb{R}$. Clearly, the set $\widetilde{C}_{(T,x_0)}$ contains some open ball. By applying the identity theorem in each variable, we can conclude that the mappings v_{eu,μ^*} and $v_{\mathrm{eu},\nu}(\vartheta, x)$ coincide on the set $(0,T) \times \mathbb{R}$. Equation (2.88) implies that $\mathbf{T}\mu^* = \mathbf{T}\nu$ must hold on $(0,T) \times \mathbb{R}$. In Lemma 2.61 it has been shown that the operator $\mathbf{T}$ is injective on the Borel measures and therefore $\mu^* = \nu$.

Step 5: Proof of Assertion (d)
For $x \geq K$ we obviously have $v_{\mathrm{eu},\mu^*}(0,x) = g(x) = 0$. Furthermore, for any $x \in [\min_{\vartheta \in (0,T]} \widehat{x}(\vartheta), K)$ we can pick a maturity $\vartheta(x) \in (0,T]$ such that $(\vartheta(x), x)$ is located on the curve, i.e. $\widehat{x}(\vartheta(x)) = x$. Due to superreplication and Assertion (b), we find that $g(x) \leq \inf_{\vartheta \in [0,T]} v_{\mathrm{eu},\mu^*}(\vartheta, x) \leq v_{\mathrm{eu},\mu^*}(\vartheta(x), x) = g(x)$ holds true.

Step 6: Proof of Assertion (e)
Assume that the curve $\widehat{x}$ is decreasing at some point. Owing to Assertion (a), we can choose some $0 < \vartheta_0 < \vartheta_1 < \vartheta_2 < T$ and some $x_0 < K$ such that $\widehat{x}(\vartheta_0) = \widehat{x}(\vartheta_2) = x_0$ and $\widehat{x}(\vartheta_1) < x_0$. From the proof of Assertion (c), we know that $\widetilde{C}_{(T,x_0)}$ is a connected subset of $C_{(T,x_0)}$. In particular, we find that (ϑ_1, x_0) is located within the continuation set. By virtue of the Assertion (b), we can therefore conclude that $g(x_0) = v_{\mathrm{am},g}(\vartheta_0, x_0) < v_{\mathrm{am},g}(\vartheta_1, x_0) \leq v_{\mathrm{am},g}(\vartheta_2, x_0) = g(x_0)$. This is impossible and hence the mapping $\vartheta \mapsto \widehat{x}(\vartheta)$ must be non-decreasing.
Now assume that there exists some $0 < \vartheta_0 < \vartheta_1 \leq T$ such that $\widehat{x}$ is constant on the interval $(\vartheta_0, \vartheta_1)$. As the curve $\widehat{x}$ is analytic, the identity theorem implies that $\widehat{x}$ is constant on $(0,T]$. In light of the second assumption, we find that $K > \widehat{x}(\vartheta) = \liminf_{\vartheta' \to 0} \widehat{x}(\vartheta') = K$ holds true for any $\vartheta \in (0,T]$ which is clearly not possible. Therefore $\vartheta \mapsto \widehat{x}(\vartheta)$ must indeed be increasing. The remaining statements now follow easily from combining all the previous assertions. □

2.4.1 A comment on the analyticity of the exercise curve

The first and the second assumption of Theorem 2.52 ensure that the curve $\vartheta \mapsto \widehat{x}(\vartheta)$ is continuous, cf. *Step 2* of Section 2.4. As demonstrated above, the analyticity of $\widehat{x}$ can be obtained by applying the implicit function theorem to the mapping $\partial_x \Psi$ where Ψ is defined as in (2.90). Equation (2.91) shows that $\partial_{xx}\Psi(\vartheta, \widehat{x}(\vartheta)) = H(\vartheta, \widehat{x}(\vartheta))$ holds true for any $\vartheta \in (0,T]$. Owing to the third assumption of Theorem 2.52, we have $H(\vartheta, \widehat{x}(\vartheta)) > 0$ for any $\vartheta \in (0,T]$ and therefore Theorem 5.7 is indeed applicable.

In this subsection we want to provide a bit of intuition under which circumstances a violation of Assumption 3 is possible. Suppose that the first and second assumption of Theorem 2.52 are satisfied. Moreover, let us assume that $c''(x) \geq 0$ for all $x < K$ where c denotes the function from (2.44). Let us remark that for the American put payoff we have $c(x) = -\frac{2r}{\sigma^2}e^K$ and therefore $c'(x) = c''(x) = 0$. We switch to forward time and write with a slight abuse of notation

$$\Psi(t,x) := v_{\mathrm{eu},\mu^*}(T-t,x) - g(x).$$

Assume there exists some $t_0 \in [0,T)$ such that

$$(\partial_{xx}\Psi)(t_0,x_0) = H(T-t_0,x_0) = 0$$

where $x_0 := \widehat{x}(T-t_0)$. To put differently, at the point (t_0,x_0) Theorem 5.7 is not straightforwardly applicable in order to warrant the analyticity of the curve $\widehat{x}$ on an open complex domain containing t_0. Owing to Lemma 5.1, we can pick for each $t \in [0,T)$ a constant $\varepsilon_t > 0$ such that $(\partial_{xx}\Psi)(t,x) \geq 0$ holds true for any $x \in \mathbb{R}$ with $|\,\widehat{x}(T-t) - x\,| \leq \varepsilon_t$. Let us *assume* that this property is locally satisfied in a uniform manner, in the sense that there exists some $t_1 \in (t_0,T)$ and some $\varepsilon > 0$ such that $(\partial_{xx}\Psi)(t,x) \geq 0$ for all $(t,x) \in \Lambda$ where

$$\Lambda := \{(t, \widehat{x}(T-t)+\delta) \mid t \in [t_0,t_1] \text{ and } |\delta| \leq \varepsilon\}.$$

Assumption 2 warrants that $\widehat{x}(T-t) < K$ for any $t \in [t_0,t_1]$. Hence, by choosing ε sufficiently small, we can always achieve that the set Λ is contained in $[t_0,t_1] \times (-\infty,K)$. Furthermore, Kolmogorov's backward equation yields

$$(\partial_t + \mathcal{A})\partial_{xx}\Psi = -\partial_{xx}\mathcal{A}g = -\frac{\sigma^2}{2}c'' \leq 0$$

on $[0,T) \times (-\infty,K)$ where $\mathcal{A} := \left(r - \frac{\sigma^2}{2}\right)\partial_x + \frac{\sigma^2}{2}\partial_{xx} - r$. Let τ denote the first exit time of the space-time process started at (t_0,x_0) from the set Λ . For any $t' \in (t_0,t_1)$ we write $\tau_{t'} := t' \wedge \tau$ and conclude by applying Dynkin's formula that

$$\begin{aligned}
0 &\leq \mathbb{E}_{(t_0,x_0)}\left[e^{-r\tau_{t'}}\partial_{xx}\Psi(\tau_{t'},X_{\tau_{t'}})\right] \\
&= e^{-rt_0}\partial_{xx}\Psi(t_0,x_0) + \mathbb{E}_{(t_0,x_0)}\left[\int_0^{\tau_{t'}} e^{-rs}((\partial_t+\mathcal{A})\partial_{xx}\Psi)(s,X_s)\,\mathrm{d}s\right] \\
&= -\frac{\sigma^2}{2}\mathbb{E}_{(t_0,x_0)}\left[\int_0^{\tau_{t'}} e^{-rs}c''(X_s)\,\mathrm{d}s\right] \leq 0
\end{aligned} \tag{2.94}$$

holds true, cf. [RW, page 254]. Due to the fact that the quantity $\partial_{xx}\Psi$ is non-negative on the set Λ, we have that $(\partial_{xx}\Psi)(\tau_{t'},X_{\tau_{t'}}) \geq 0$. As $t' \in (t_0,t_1)$ was chosen arbitrarily, Equation (2.94) implies that $\partial_{xx}\Psi = 0$ on the interior of the set Λ. The analyticity of the function $\partial_{xx}\Psi$ in both components now yields that $\partial_{xx}\Psi$ vanishes on $(0,T) \times (-\infty,K)$. Applying the fundamental theorem of calculus, we obtain

$$(\partial_x\Psi)(t,x) = (\partial_x\Psi)(t,\widehat{x}(T-t)) = 0$$

for any $(t,x) \in (0,T) \times (-\infty,K)$. This clearly contradicts the uniqueness of the minima imposed by Assumption 2.

Consequently, when attempting the construction of an example in our setting where $(\partial_{xx}\Psi)(t_0,x_0) = H(T-t_0,x_0) = 0$ holds true for some $t_0 \in [0,T)$, one must exclude the existence of the set Λ. Moreover, we note that the specific geometric properties of Λ are mostly insignificant. The reader may think of differently shaped sets where the latter arguments work.

3 Computational methods and numerical results

3.1 Approximate CDEOs

Suppose that $(T, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$ and let $g : \mathbb{R} \to \mathbb{R}_+$ denote some American payoff function. The computation of the cheapest dominating European option of g relative to T, x_0 in the sense of Section 2.2 requires us to solve the following general capacity problem

$$\begin{array}{ll} \text{minimize} & v_{\mathrm{eu},\mu}(T, x_0) \\ \text{subject to} & \mu \in \mathcal{M}^+(\mathbb{R}), \\ & v_{\mathrm{eu},\mu}(\vartheta, x) \geq g(x) \text{ for any } (\vartheta, x) \in (0, T) \times \mathbb{R} \end{array} \tag{GCAP}$$

where $v_{\mathrm{eu},\mu}$ is defined as in (2.47). In the Sections 2.1 and 2.2 we discussed that CDEOs provide us with natural candidates for representing European payoffs. As a quick reminder, let us recall some central assertions from Proposition 2.3 and Proposition 2.10 in a slightly generalized setting.

3.1 Lemma: *Let $g : \mathbb{R} \to \mathbb{R}_+$ denote a lower semi-continuous American payoff satisfying* (1.5)*. For any GCAP-admissible measure μ and any $(\vartheta, x) \in (0, T) \times \mathbb{R}$ we have $v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{eu},\mu}(\vartheta, x)$.*

Proof. Choose $(\vartheta, x) \in (0, T) \times \mathbb{R}$ and let τ denote some $[0, \vartheta]$-valued stopping time. The mapping g is assumed to be lower semi-continuous and consequently we have $g(X_\tau) \leq \liminf_{\vartheta' \nearrow \vartheta} g(X_{\tau \wedge \vartheta'}) \leq \liminf_{\vartheta' \nearrow \vartheta} v_{\mathrm{eu},\mu}(\vartheta - \tau \wedge \vartheta', X_{\tau \wedge \vartheta'})$. Fatou's lemma and the martingale property established in (2.93) imply

$$\mathbb{E}_x \left[e^{-r\tau} g(X_\tau) \right] \leq \liminf_{\vartheta' \nearrow \vartheta} \mathbb{E}_x \left[e^{-r(\tau \wedge \vartheta')} v_{\mathrm{eu},\mu}(\vartheta - \tau \wedge \vartheta', X_{\tau \wedge \vartheta'}) \right] = v_{\mathrm{eu},\mu}(\vartheta, x).$$

Maximizing the left-hand side over all $[0, \vartheta]$-valued stopping proves the assertion. □

3.2 Lemma: *Let $g : \mathbb{R} \to \mathbb{R}_+$ be a lower semi-continuous American payoff satisfying* (1.5)*. Suppose that $(T, x_0) \in [0, T] \times \mathbb{R}$ is located within the continuation set associated to g and that $C_{(T,x_0)}$, as defined in Subsection 2.1.1, is relatively open in $[0, T] \times \mathbb{R}$. Furthermore, assume there exists a measure $\mu^* \in \mathcal{M}^+(\mathbb{R})$ such that μ^* represents g with respect to (T, x_0), i.e. $v_{\mathrm{am},g}(\vartheta, x) \leq v_{\mathrm{eu},\mu^*}(\vartheta, x)$ for any $(\vartheta, x) \in (0, T) \times \mathbb{R}$ and equality holds on $C_{(T,x_0)}$. Then:*

1. *The representing measure μ^* is unique.*

2. *The measure μ^* is the unique solution of Program GCAP.*

Proof.

1. Assume there exists another measure $\tilde{\mu}$ which represents g with respect to (T, x_0). Clearly, the mappings v_{eu,μ^*} and $v_{\mathrm{eu},\tilde{\mu}}$ coincide on some non-empty open set $O \subset C_{(T,x_0)}$. Lemma 2.61 implies that $\mu^* = \tilde{\mu}$.

2. We have $e^{-rT}\|\mu^*\| = v_{\mathrm{eu},\mu^*}(T, x_0) = v_{\mathrm{am},g}(T, x_0)$ and therefore μ^* is GCAP-optimal. Hence, any other GCAP-optimal measure ν^* must satisfy $v_{\mathrm{eu},\nu^*}(T, x_0) = v_{\mathrm{am},g}(T, x_0)$. In light of *Step 1* from Section 2.4, we find that the mappings v_{eu,μ^*} and v_{eu,ν^*} are analytic on an open $\mathbb{C}^2$-domain containing the set $(0, T) \times \mathbb{R}$. We can choose some radius $0 < \delta < \frac{T}{2}$ such that

$$D_\delta := \{(\vartheta, x) \in (0, T) \times \mathbb{R} \mid \|(\vartheta, x) - (T, x_0)\|_2 \leq \delta\}$$

is a subset of $C_{(T,x_0)}$. Let τ_δ denote the first exit time of the space-time process started at time 0 and initial log-price x_0 from the set D_δ. Due to the fact that μ^* represents g with respect to (T, x_0), we find that $v_{\mathrm{eu},\nu^*}(T - \tau_\delta, X_{\tau_\delta}) \geq v_{\mathrm{am},g}(T - \tau_\delta, X_{\tau_\delta}) = v_{\mathrm{eu},\mu^*}(T - \tau_\delta, X_{\tau_\delta})$ holds true $\mathbb{P}_{x_0}$-almost surely. The stopping time τ_δ is $(0, \delta]$-valued and taking (2.93) into account, we obtain by optional sampling

$$\mathbb{E}_{x_0}\left[\left(\hat{v}_{\mathrm{eu},\nu^*} - \hat{v}_{\mathrm{eu},\mu^*}\right)(T - \tau_\delta, X_{\tau_\delta})\right] = \hat{v}_{\mathrm{eu},\nu^*}(T, x_0) - \hat{v}_{\mathrm{am},g}(T, x_0) = 0 \tag{3.3}$$

where $\hat{v}_{\mathrm{eu},\mu}(T - t, x) := e^{-rt} v_{\mathrm{eu},\mu}(T - t, x)$ denotes the discounted European value at time t and log-price x. As a matter of fact, Equation (3.3) holds true for any $\delta' \in (0, \delta)$ which implies that the mappings $\hat{v}_{\mathrm{eu},\nu^*}$ and $\hat{v}_{\mathrm{eu},\mu^*}$ coincide on the set D_δ. Owing to Lemma 2.61 and (2.88), we can finally conclude that $\mu^* = \nu^*$.

□

In case that the payoff g satisfies the regularity conditions from Section 2.2, Theorem 2.51 warrants that the optimal value in GCAP is obtained by some admissible measure. Let us remark that [LW, Theorem 2.1] is not straightforwardly applicable in order to verify the existence of an optimizer. This is mainly due to the lacking compactness of the underlying spaces and the possibly singular behavior of the mapping $v_{\mathrm{eu},\mu}$ at the time boundaries. Within the scope of this thesis, we will not further discuss any aspects related to the existence of optimizers in the rather general optimization task GCAP. Instead we focus on certain semi-infinite linear programs which serve as numerically feasible surrogates for GCAP.

We sketch a simple way how to obtain approximate solutions of GCAP numerically. To this end we adopt the following well-known discretization approach, cf. [CR], [HK] and [LW]. The cone $\mathcal{M}^+(\mathbb{R})$ appearing in GCAP is replaced by some suitable finite dimensional subcone. This yields an approximation to GCAP in the form of a linear program

P_n with a finite number of variables and an infinite number of constraints. Optimization tasks of this type fall within the scope of semi-infinite programming theory. First, we want to prove the following existence and consistency result. In order to simplify the line of argument, it is convenient to include a certain measure which is related to an invariant function of the Black-Scholes pricing semi-group into the discretization.

3.4 Proposition: *Define* $f_1(y) := e^{-\frac{2r}{\sigma^2}y} + e^y$ *and* $\mathrm{d}\mu_1(y) := \mathrm{N}(x_0 + \hat{r}T, \sigma^2 T, y)\, f_1(y)\mathrm{d}y$. *Suppose that* $\mu_2, \mu_3, ... \in \mathcal{M}(\mathbb{R})$ *and write* $U_n := \{\sum_{k=1}^n a_k \mu_k \mid a \in \mathbb{R}^n\}$. *Moreover, we denote by* U_∞ *the closure of the set* $\cup_{n\in\mathbb{N}} U_n$ *with respect to the vague topology. Suppose that* $g : \mathbb{R} \to \mathbb{R}_+$ *is a continuous American payoff satisfying* $\|\frac{g}{f_1}\|_\infty < \infty$. *Then:*

1. *For any* $n \in \mathbb{N} \cup \{\infty\}$ *the optimal value* p_n *in the linear program*

$$\begin{aligned} &\text{minimize} && v_{\mathrm{eu},\mu}(T, x_0) \\ &\text{subject to} && \mu \in U_n \cap \mathcal{M}^+(\mathbb{R}), \\ & && v_{\mathrm{eu},\mu}(\vartheta, x) \geq g(x) \text{ for any } (\vartheta, x) \in (0,T) \times \mathbb{R} \end{aligned} \tag{P_n}$$

   is attained by some admissible measure μ_n^* *satisfying* $p_n = e^{-rT}\|\mu_n^*\| \leq f_1(x_0)\|\frac{g}{f_1}\|_\infty$. *For any* $m \leq n$ *the measure* μ_m^* *is* P_n*-admissible and we have* $p_m \geq p_n$.

2. *A subsequence of optimizers* $\mu_{n_k}^*$ *converges vaguely to some* P_∞*-admissible measure* ν_∞. *Moreover, the optimal values satisfy the inequality*

$$p_\infty \leq v_{\mathrm{eu},\nu_\infty}(T, x_0) \leq \inf_{n\in\mathbb{N}} p_n = \lim_{n\to\infty} p_n.$$

3. *In addition, suppose that there exists some sequence* $\xi_n \in U_n \cap \mathcal{M}^+(\mathbb{R})$ *which converges weakly to a* P_∞*-optimal measure* μ_∞^* *as* $n \to \infty$. *If*

$$\lim_{n\to\infty} \sup_{(\vartheta,x)\in(0,T)\times\mathbb{R}} \frac{|v_{\mathrm{eu},\mu_\infty^* - \xi_n}(\vartheta, x)|}{f_1(x)} = 0$$

   holds true, we can conclude that the measure ν_∞ *from Assertion 2 is* P_∞*-optimal and that the subsequence* $\mu_{n_k}^*$ *converges weakly to* ν_∞. *In particular, we have*

$$p_\infty = v_{\mathrm{eu},\nu_\infty}(T, x_0) = \inf_{n\in\mathbb{N}} p_n = \lim_{n\to\infty} p_n.$$

Remark: Replacing U_n *by* $\widetilde{U}_n := \{\sum_{k=1}^n a_k \mu_k \mid a \in \mathbb{R}_+^n\}$ *does not effect the validity of the latter proposition.*

Proof.

1. For any $n \in \mathbb{N} \cup \{\infty\}$ we consider instead of P_n the equivalent program

$$\begin{aligned} &\text{minimize} && e^{-rT}\|\mu\| \\ &\text{subject to} && \mu \in U_n \cap \mathcal{M}^+(\mathbb{R}), \\ & && \int_{-\infty}^{\infty} \widetilde{\kappa}(\vartheta, x, y)\mathrm{d}\mu(y) \geq \widetilde{g}(x) \text{ for any } (\vartheta, x) \in \Lambda \end{aligned} \tag{P'_n}$$

where $\Lambda := (0,T) \times \mathbb{R}$, $\widetilde{g} := \frac{g}{f_1} \in C_b(\mathbb{R})$ and

$$\widetilde{\kappa}(\vartheta, x, y) := e^{-r\vartheta} \frac{\mathrm{N}\left(x + \hat{r}\vartheta, \sigma^2\vartheta, y\right)}{f_1(x)\mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right)}.$$

The set of all P'_n-admissible measures is denoted by $Z_{\mathrm{P}'_n}$. Choose $n \in \mathbb{N} \cup \{\infty\}$ arbitrarily. In order to show that the optimal value in Program P'_n is attained, define the measure $\mu_{\mathrm{a}} := \|\widetilde{g}\|_\infty \, \mu_1 \in U_n \cap \mathcal{M}^+(\mathbb{R})$. Due to the fact that f_1 is an invariant function of the Black-Scholes pricing semi-group, we obtain

$$\int_{-\infty}^{\infty} \widetilde{\kappa}(\vartheta, x, y) \mathrm{d}\mu_{\mathrm{a}}(y) = \|\widetilde{g}\|_\infty \frac{e^{-r\vartheta}\mathbb{E}_x[f_1(X_\vartheta)]}{f_1(x)} = \|\widetilde{g}\|_\infty \geq \widetilde{g}(x)$$

for any $(\vartheta, x) \in \Lambda$. This shows that μ_{a} is P'_n-admissible which indicates that any potential minimizer $\mu^* \in Z_{\mathrm{P}'_n}$ must satisfy the inequality

$$\|\mu^*\| \leq \|\mu_{\mathrm{a}}\| = e^{rT} \|\widetilde{g}\|_\infty f_1(x_0) =: \rho.$$

In other words, the measure μ^* must be contained in the vaguely compact ball $B_{\mathcal{M}}(\rho) = \{\mu \in \mathcal{M}(\mathbb{R}) \mid \|\mu\| \leq \rho\}$. In Program P'_n it is therefore sufficient to minimize over set

$$Z_{\mathrm{P}'_n} \cap B_{\mathcal{M}}(\rho) = \bigcap_{(\vartheta,x)\in\Lambda} H(\vartheta, x) \,\cap\, U_n \,\cap\, \mathcal{M}^+(\mathbb{R}) \,\cap\, B_{\mathcal{M}}(\rho)$$

where $H(\vartheta, x) := \{\mu \in \mathcal{M}(\mathbb{R}) \mid \int_{-\infty}^{\infty} \widetilde{\kappa}(\vartheta, x, y) \mathrm{d}\mu(y) \geq \widetilde{g}(x)\}$. For any $(\vartheta, x) \in \Lambda$ the mapping $\mathbb{R} \ni y \mapsto \widetilde{\kappa}(\vartheta, x, y)$ vanishes at infinity and therefore the associated half-space $H(\vartheta, x)$ is vaguely closed. Due to the fact that the cone $U_n \cap \mathcal{M}^+(\mathbb{R})$ is vaguely closed as well, we can conclude that $Z_{\mathrm{P}'_n} \cap B_{\mathcal{M}}(\rho)$ is a vaguely compact subset of $\mathcal{M}(\mathbb{R})$. Theorem 5.18 implies that the target functional $\mu \mapsto e^{-rT}\|\mu\|$ is lower semi-continuous with respect to the vague topology. Therefore, the optimal value p_n in Program P'_n is attained by some measure $\mu_n^* \in Z_{\mathrm{P}'_n} \cap B_{\mathcal{M}}(\rho)$, cf. Lemma 5.9.

2. From Assertion 1 we know that $\|\mu_n^*\| \leq \rho$ for any $n \in \mathbb{N}$. General theory tells us that the vague topology is metrizable on the total variation unit ball and Theorem 5.18 warrants that the latter set is vaguely compact. Hence, there exists a subsequence $\mu_{n_k}^*$ which converges vaguely to some measure $\nu_\infty \in B_{\mathcal{M}}(\rho) \cap \mathcal{M}^+(\mathbb{R})$. For any $(\vartheta, x) \in \Lambda$ the mapping $\mathbb{R} \ni y \mapsto \widetilde{\kappa}(\vartheta, x, y)$ vanishes at infinity which shows that
$$\int_{-\infty}^{\infty} \widetilde{\kappa}(\vartheta, x, y) \mathrm{d}\nu_\infty(y) = \lim_{k\to\infty} \int_{-\infty}^{\infty} \widetilde{\kappa}(\vartheta, x, y) \mathrm{d}\mu_{n_k}^*(y) \geq \widetilde{g}(x).$$
In other words, the measure ν_∞ is admissible in Program P_∞ and therefore $p_\infty \leq v_{\mathrm{eu},\nu_\infty}(T, x_0)$. By vague convergence we find that $\|\nu_\infty\| \leq \liminf_{k\to\infty} \|\mu_{n_k}^*\|$, cf. [KL, Lemma 13.15]. Due to the monotonicity of the P_n-optimal values, we can finally conclude that $p_\infty \leq v_{\mathrm{eu},\nu_\infty}(T, x_0) \leq \inf_{n\in\mathbb{N}} p_n = \lim_{n\to\infty} p_n$ holds true, as claimed.

3. For any $n \in \mathbb{N}$ we define the non-negative measure $\eta_n := \xi_n + \varepsilon_n \mu_1$ where
$$\varepsilon_n := \sup_{(\vartheta,x)\in(0,T)\times\mathbb{R}} \frac{|v_{\mathrm{eu},\mu^*_\infty-\xi_n}(\vartheta,x)|}{f_1(x)}.$$
Clearly, the measure μ^*_∞ is P_∞-admissible which yields that
$$\begin{aligned} v_{\mathrm{eu},\eta_n}(\vartheta,x) - g(x) &\geq v_{\mathrm{eu},\eta_n}(\vartheta,x) - v_{\mathrm{eu},\mu^*_\infty}(\vartheta,x) \\ &= f_1(x)\left(\varepsilon_n - \frac{v_{\mathrm{eu},\mu^*_\infty-\xi_n}(\vartheta,x)}{f_1(x)}\right) \geq 0 \end{aligned} \tag{3.5}$$
holds true for any $(\vartheta,x) \in \Lambda$. Consequently, the measure η_n is admissible in Program P_n and owing to the second assertion we obtain
$$p_\infty \leq v_{\mathrm{eu},\nu_\infty}(T,x_0) \leq p_n \leq v_{\mathrm{eu},\eta_n}(T,x_0) = v_{\mathrm{eu},\xi_n}(T,x_0) + \varepsilon_n f_1(x_0).$$
As the sequence ξ_n was assumed to converge weakly to the measure μ^*_∞, we find that $\lim_{n\to\infty} v_{\mathrm{eu},\xi_n}(T,x_0) = v_{\mathrm{eu},\mu^*_\infty}(T,x_0) = p_\infty$ and this finally shows that
$$p_\infty = v_{\mathrm{eu},\nu_\infty}(T,x_0) = \lim_{n\to\infty} p_n = \lim_{n\to\infty} v_{\mathrm{eu},\eta_n}(T,x_0) = \lim_{n\to\infty} v_{\mathrm{eu},\xi_n}(T,x_0).$$
In particular, we obtain $\lim_{n\to\infty} \|\mu^*_n\| = \|\nu_\infty\|$. In Assertion 2 it was shown that some subsequence of the optimizers, say $\mu^*_{n_k}$, converges vaguely to ν_∞. As $\mathbb{R}$ is locally compact and Polish with respect to the Euclidean topology, the Portmanteau theorem warrants that $\mu^*_{n_k}$ converges weakly to the measure ν_∞, cf. [KL, Theorem 13.16].

□

Within the scope of optimal stopping theory, one of the fundamental technical assumptions about the American payoff of interest is the integrability condition (1.5). Conveniently, the growth condition $\|\frac{g}{f_1}\|_\infty < \infty$ from Proposition 3.4 warrants that (1.5) holds true.

3.6 Lemma: *Let $f_1 : \mathbb{R} \to \mathbb{R}_+$ denote a non-negative invariant function of the Black-Scholes pricing semi-group, e.g. $f_1(x) := e^{-\frac{2r}{\sigma^2}x} + e^x$. Any continuous American payoff $g : \mathbb{R} \to \mathbb{R}_+$ satisfying $\|\frac{g}{f_1}\|_\infty < \infty$ satisfies the integrability condition* (1.5) *as well.*

Proof. Clearly, there exists some constant $C > 0$ such that $g \leq Cf_1$. We assumed that $v_{\mathrm{eu},f_1} = f_1$ holds true on $\mathbb{R}_+ \times \mathbb{R}$. Applying Kolmogorov's backward equation yields $\mathcal{A}f_1 = 0$ where $\mathcal{A} := (r - \frac{\sigma^2}{2})\partial_x + \frac{\sigma^2}{2}\partial_{xx} - r$ denotes the generator of the pricing semi-group on the twice differentiable functions. By virtue of Ito's formula, we obtain
$$e^{-rt}g(X_t) \leq Ce^{-rt}f_1(X_t) \leq Cf_1(X_0) + C\sigma\left|\int_0^t e^{-rs}f_1'(X_s)\mathrm{d}W_s\right|.$$
The process $V_t := \left|\int_0^t e^{-rs}f_1'(X_s)\mathrm{d}W_s\right|$ is a non-negative sub-martingale with continuous paths. Applying Doob's L_2-inequality shows that
$$\left(\mathbb{E}_x \sup_{t\in[0,T]} V_t\right)^2 \leq \mathbb{E}_x\left[\left(\sup_{t\in[0,T]} V_t\right)^2\right] \leq 4\,\mathbb{E}_x\left[V_T^2\right]$$

for any $x \in \mathbb{R}$. In light of the Ito-isometry, we find that

$$\mathbb{E}_x\left[V_T^2\right] = \int_0^T e^{-2rs}\,\mathbb{E}_x\left[(f_1'(X_s))^2\right]\mathrm{d}s =: \beta^2(x) < \infty.$$

Let us remark that the quantity $\mathbb{E}_x\left[(f_1'(X_s))^2\right]$ can be written as a linear combination of characteristic functions associated to X_s. Hence, the constant $\beta(x)$ can be explicitly calculated. In conclusion, we have shown that

$$\mathbb{E}_x\left[\sup_{t\in[0,T]} e^{-rt} g(X_t)\right] \leq C f_1(x) + 2C\sigma\beta(x) < \infty$$

holds true for any $x \in \mathbb{R}$. Hence, the American payoff g satisfies condition (1.5). □

In case that the American payoff is represented by some measure, we obtain the following simple uniqueness result.

3.7 Lemma: *Write $f_1(y) := e^{-\frac{2r}{\sigma^2}y} + e^y$ and $\mathrm{d}\mu_1(y) := \mathrm{N}(x_0 + \hat{r}T, \sigma^2 T, y)\, f_1(y)\mathrm{d}y$. Suppose that $\mu_2, \mu_3, ... \in \mathcal{M}(\mathbb{R})$ and let U_n be defined as in Proposition 3.4. Let $g : \mathbb{R} \to \mathbb{R}_+$ denote a continuous American payoff satisfying $\|\frac{g}{f_1}\|_\infty < \infty$ which is represented relative to (T, x_0) by some measure $\mu^* \in \mathcal{M}^+(\mathbb{R})$. Assume there exists some $N \in \mathbb{N}\cup\{\infty\}$ such that $\mu^* \in U_N$. Then:*

1. *For any $n \in \{N, .., \infty\}$ the measure μ^* is the unique solution of Program P_n.*
2. *If the assumptions of Assertion 3 from Proposition 3.4 are satisfied, there exists a subsequence of P_{n_k}-optimal measures $\mu^*_{n_k}$ which converges weakly to μ^*. In particular, we have $v_{\mathrm{eu},\mu^*}(\vartheta, x) = \lim_{k\to\infty} v_{\mathrm{eu},\mu^*_{n_k}}(\vartheta, x)$ for any $(\vartheta, x) \in (0, T) \times \mathbb{R}$.*

Proof.

1. Lemma 3.6 warrants that the payoff g satisfies (1.5). For $n \in \{N, .., \infty\}$ let μ^*_n denote a solution of P_n. Clearly, we have $v_{\mathrm{eu},\mu^*}(T, x_0) = v_{\mathrm{eu},\mu^*_n}(T, x_0)$ and therefore μ^*_n solves GCAP. From Lemma 3.2 we know that μ^* is the unique solution of Program GCAP and this implies the validity of the assertion.
2. If $N < \infty$, the assertion at hand is trivially satisfied as we have $\mu^*_n = \mu^*$ for any $n \geq N$. In case that $N = \infty$, we know from Proposition 3.4 that a subsequence $\mu^*_{n_k}$ converges weakly to some P_∞-optimal measure ν_∞. Owing to Assertion 1, the measures ν_∞ and μ^* coincide. The other claim follows from the fact that for any $(\vartheta, x) \in (0, T) \times \mathbb{R}$ the integral kernel occurring in (2.47) vanishes at infinity.

□

In our numerical experiments it was convenient to consider non-negative, absolutely continuous discretizations in Proposition 3.4. To be more specific, we chose measures $\mu_1, ..., \mu_n \in \mathcal{M}^+(\mathbb{R})$ which are absolutely continuous with respect to the Gaussian law

$\mathcal{N}(x_0 + \hat{r}T, \sigma^2 T)$, that is to say $\mathrm{d}\mu_k(y) := f_k(y)\mathrm{N}(x_0 + \hat{r}T, \sigma^2 T, y)\,\mathrm{d}y$ for some non-negative function f_k. In this case, Program P_n with $\widetilde{U}_n$ from Proposition 3.4 can be equivalently rewritten as

$$\begin{aligned} &\text{minimize} && V(T, x_0)^\top a \\ &\text{subject to} && a \in \mathbb{R}^n_+, \\ &&& V(\vartheta, x)^\top a \geq g(x) \text{ for any } (\vartheta, x) \in [0, T] \times \mathbb{R} \end{aligned} \tag{$\mathrm{P}^{\mathrm{c}}_n$}$$

where $V(\vartheta, x) := (v_{\mathrm{eu},f_1}(\vartheta, x), ..., v_{\mathrm{eu},f_n}(\vartheta, x)) \in \mathbb{R}^n$ and v_{eu,f_k} is defined as in (2.45). The Lagrange dual problem associated to $\mathrm{P}^{\mathrm{c}}_n$ is given by

$$\begin{aligned} &\text{maximize} && \langle g, \lambda \rangle \\ &\text{subject to} && \lambda \in \mathcal{M}^+([0, T] \times \mathbb{R}), \\ &&& \langle V, \lambda \rangle \leq V(T, x_0). \end{aligned} \tag{$\mathrm{D}^{\mathrm{c}}_n$}$$

Here we denote by $\langle \cdot, \lambda \rangle := \int_{[0,T]\times\mathbb{R}} \cdot \,\mathrm{d}\lambda$ the $\mathbb{R}^n$-valued integral with respect to the measure λ. Imposing some additional regularity assumptions warrants that the dual optimal value is always attained and that we have strong duality.

3.8 Lemma: *Define $f_1(x) := e^{-\frac{2r}{\sigma^2}x} + e^x$ and suppose that $f_2, f_3, ... : \mathbb{R} \to \mathbb{R}_+$ are continuous functions such that $f_1, f_2, ...$ are linearly independent, $v_{\mathrm{eu},f_k}(T + \varepsilon, x_0) < \infty$ for some $\varepsilon > 0$ and*

$$\lim_{y \to \infty} \sup_{\substack{\vartheta \in [0,T] \\ |x| > y}} \frac{v_{\mathrm{eu},f_k}(\vartheta, x)}{f_1(x)} = 0 \tag{3.9}$$

for any $k \in \{2, 3, ...\}$. Moreover, let $g : \mathbb{R} \to \mathbb{R}_+$ denote a continuous American payoff satisfying the growth condition

$$\lim_{|x| \to \infty} \frac{g(x)}{f_1(x)} = 0 \tag{3.10}$$

and choose $n \in \mathbb{N}$ arbitrarily. Then:

1. *The P^c_n-optimal value p_n is attained by some admissible vector $a^{*,n}$ and the optimal value d_n in D^c_n is attained by some admissible measure λ^*_n.*

   The associated P_n-optimal measure $\mu^*_n = \sum_{k=1}^n a^{*,n}_k \mu_k$ satisfies $\|\mu^*_n\| \leq e^{rT} f_1(x_0) \|\frac{g}{f_1}\|_\infty$ and we have $\|\lambda^*_n\| \leq f(x_0)$.

2. *We have*

$$g(x_0) \leq d_n = p_n \leq f_1(x_0) \sup_{x \in \mathbb{R}} \frac{g(x)}{f_1(x)}$$

   and the complementary slackness equations

$$\begin{aligned} V(\vartheta, x)^\top a^{*,n} &= g(x) && \lambda^*_n\text{-a.e. on } [0, T] \times \mathbb{R}, \\ \langle V, \lambda^*_n \rangle_k &= V(T, x_0)_k && \text{if } a^{*,n}_k > 0 \end{aligned} \tag{3.11}$$

   hold true.

3. *A subsequence of dual optimizers $(\lambda^*_{n_k})_{k\in\mathbb{N}}$ converges vaguely to some measure λ^*_∞ which is D^c_n-admissible for any $n \in \mathbb{N}$. Moreover, we have*

$$\lim_{n\to\infty} p_n = \lim_{n\to\infty} d_n = \langle g, \lambda^*_\infty \rangle.$$

4. *The D^c_n-optimizer λ^*_n is discrete and concentrated on not more than $n+1$ points of the set $[0,T]\times\mathbb{R}$.*

Proof. We write $\Lambda := [0,T]\times\mathbb{R}$. Instead of P^c_n, we can consider the equivalent program

$$\begin{aligned} &\text{minimize} && V(T,x_0)^\top a \\ &\text{subject to} && a \in \mathbb{R}^n_+, \\ &&& \widetilde{V}^\top a \geq \widetilde{g} \text{ on } \Lambda \end{aligned} \tag{$\mathrm{P}^{\mathrm{c}\,\prime}_n$}$$

and its Lagrange dual

$$\begin{aligned} &\text{maximize} && \langle \widetilde{g}, \lambda \rangle \\ &\text{subject to} && \lambda \in \mathcal{M}^+(\Lambda), \\ &&& \langle \widetilde{V}, \lambda \rangle \leq V(T,x_0) \end{aligned} \tag{$\mathrm{D}^{\mathrm{c}\,\prime}_n$}$$

where $\widetilde{V} := \frac{V}{f_1}$ and $\widetilde{g} := \frac{g}{f_1}$. Due to the growth condition (3.10), we find that $\widetilde{g}$ lives in $C_0(\Lambda)$ and therefore $\lambda \mapsto \langle \widetilde{g}, \lambda \rangle$ constitutes a vaguely continuous linear functional on the space $\mathcal{M}(\Lambda)$.

1. Clearly, the mapping $\mathbb{R}^n \ni a \mapsto \sum_{k=1}^n a_k \mu_k$ constitutes an isomorphism between $\mathbb{R}^n$ and the span of $\{\mu_1, ..., \mu_n\}$. By virtue of Proposition 3.4, we obtain that the optimal value p_n in $\mathrm{P}^{\mathrm{c}\,\prime}_n$ is attained by some admissible vector $a^{*,n} \in \mathbb{R}^n_+$ satisfying $\| \sum_{k=1}^n a^{*,n}_k \mu_k \| = e^{rT} p_n \leq e^{rT} f_1(x_0) \|\widetilde{g}\|_\infty$.

   Next, we show that the optimal value in Program $\mathrm{D}^{\mathrm{c}\,\prime}_n$ is attained. Clearly, the point mass $f_1(x_0)\delta_{(T,x_0)}$ is contained in the $\mathrm{D}^{\mathrm{c}\,\prime}_n$-admissible set $Z_{\mathrm{D}^{\mathrm{c}\,\prime}_n}$ and we have $d_n \geq \langle \widetilde{g}, f_1(x_0)\delta_{(T,x_0)} \rangle = g(x_0)$. Due to the fact that f_1 is an invariant function of the Black-Scholes pricing semi-group, we have $\widetilde{V}_1 = 1$ which implies that

$$Z_{\mathrm{D}^{\mathrm{c}\,\prime}_n} = \bigcap_{k=2}^n H_k \,\cap\, B_{\mathcal{M}(\Lambda)}(f_1(x_0)) \,\cap\, \mathcal{M}^+(\Lambda) \tag{3.12}$$

   where $H_k := \{\lambda \in \mathcal{M}(\Lambda) \mid \langle \widetilde{V}_k, \lambda \rangle \leq v_{\mathrm{eu},f_k}(T,x_0)\}$ and $B_{\mathcal{M}(\Lambda)}(f_1(x_0))$ denotes the closed total variation ball of radius $f_1(x_0)$. From Theorem 5.18 we know that the latter set is compact with respect to the vague topology on $\mathcal{M}(\Lambda)$. Moreover, the growth condition (3.9) warrants that $\widetilde{V}_2, ..., \widetilde{V}_n \in C_0(\Lambda)$ and consequently the half-spaces $H_2, ..., H_n$ are vaguely closed. The same holds true for the cone $\mathcal{M}^+(\Lambda)$ and therefore we can finally conclude that $Z_{\mathrm{D}^{\mathrm{c}\,\prime}_n}$ is a vaguely compact subset of the space $\mathcal{M}(\Lambda)$. As argued above, the target functional $\lambda \mapsto \langle \widetilde{g}, \lambda \rangle$ is vaguely continuous and this implies that the $\mathrm{D}^{\mathrm{c}\,\prime}_n$-optimal value $d_n \geq g(x_0)$ is attained by some admissible measure λ^*_n satisfying $\|\lambda^*_n\| \leq f_1(x_0)$.

2. The strong duality $p_n = d_n$ will be established in a similar fashion as in the proof of Theorem 2.63. Most of the arguments below strongly resemble their counterparts from Subsection 2.3.3. Thus, we dare to skip a few simple calculations in order to shorten the exposition below. For the optimization task at hand, we consider the Lagrange function
$$K : \mathbb{R}^n \times \mathcal{M}(\Lambda) \times \mathbb{R}^n \to [-\infty, \infty]$$
which is defined by
$$K(a, \lambda, \xi) := V(T, x_0)^\top a + \langle \tilde{g}, \lambda \rangle - \langle \widetilde{V}, \lambda \rangle^\top a - \xi^\top a - \mathcal{I}_{\mathcal{M}^+(\Lambda)\times\mathbb{R}^n_+}(\lambda, \xi). \quad (3.13)$$
Individually analyzing the summands in the latter equation yields that the mapping K is linear in the variable a and concave in the variable (λ, ξ). It is easy to see that for any fixed $(\lambda, \xi) \in \mathcal{M}(\Lambda) \times \mathbb{R}^n$ the linear mapping $K_{(\lambda,\xi)}(a) := K(a, \lambda, \xi)$ is closed, c.f. Section 5.4. Analogously to Lemma 2.74, the dual value function $v : \mathbb{R}^n \to [-\infty, \infty]$ is defined by
$$v(b) := \inf_{(\lambda,\xi)\in\mathcal{M}(\Lambda)\times\mathbb{R}^n} K^*_{(\lambda,\xi)}(b) = \inf_{(\lambda,\xi)\in\mathcal{M}(\Lambda)\times\mathbb{R}^n} \sup_{a\in\mathbb{R}^n} \left\{ a^\top b - K(a, \lambda, \xi) \right\}.$$
Calculations similar to (2.72) and (2.73) show that $-p_n = v^{**}(0) \le v(0) = -d_n$ holds true. By virtue of Theorem 5.29, we have $v^{**} = \mathrm{cl}(\mathrm{co}(v))$. In order to establish the desired equality $v^{**}(0) = v(0)$, we follow the line of argumentation from the proof of Theorem 2.63. From there we know that it is sufficient to verify that v is a convex mapping which is lower semi-continuous at the origin.

The convexity of v follows from nearly the same calculation as in (2.76). The continuity at 0 will be established by means of Lemma 2.77. To this end we need to show that the mapping v never assumes the value $-\infty$ and that v is bounded from above on some open set containing the origin. Assume that there exists some $\tilde{b} \in \mathbb{R}^n$ such that $v(\tilde{b}) = -\infty$. Clearly, for any $a \in \mathbb{R}^n$ we have
$$v(\tilde{b}) \ge a^\top \tilde{b} - \sup_{(\lambda,\xi)\in\mathcal{M}(\Lambda)\times\mathbb{R}^n} K(a, \lambda, \xi)$$
and therefore $\sup_{(\lambda,\xi)\in\mathcal{M}(\Lambda)\times\mathbb{R}^n} K(a, \lambda, \xi) = \infty$. Minimizing over all $a \in \mathbb{R}^n$ yields
$$p_n = -v^{**}(0) = \inf_{a\in\mathbb{R}^n} \sup_{(\lambda,\xi)\in\mathcal{M}(\Lambda)\times\mathbb{R}^n} K(a, \lambda, \xi) = \infty$$
which clearly contradicts the first assertion of the lemma at hand. We conclude that v indeed only assumes values in $(-\infty, \infty]$. Moreover, for any $b \in \mathbb{R}^n$ such that $\|b\|_\infty < \frac{1}{2} \min_{k\in\{1,\dots,n\}} V(T, x_0) =: \gamma$ we have
$$\begin{aligned} v(b) &\le \sup_{a\in\mathbb{R}^n} \left\{ a^\top b - K(a, 0, 0) \right\} \\ &= \sup_{a\in\mathbb{R}^n} \left(b - V(T, x_0) \right)^\top a = 0. \end{aligned}$$

To put differently, we have $\sup_{b \in B_\infty(\gamma)} v(b) \leq 0$ and therefore Lemma 2.77 is applicable. The latter warrants that the mapping v is indeed continuous at 0 with respect to any norm topology on $\mathbb{R}^n$. Finally, we obtain $-p_n = v^{**}(0) = \mathrm{cl}(\mathrm{co}(v))(0) = \mathrm{lsc}(v)(0) = v(0) = -d_n$. The complementary slackness equations can now be derived by the standard argument

$$0 \leq \langle \widetilde{V}^\top a^{*,n} - \widetilde{g}, \lambda_n^* \rangle = \langle \widetilde{V}, \lambda_n^* \rangle^\top a^{*,n} - V(T, x_0)^\top a^{*,n} \leq 0.$$

3. From Assertion 1 we know that any dual optimizer satisfies $\|\lambda_n^*\| \leq f_1(x_0)$. In regard of Theorem 5.18 and the metrizability of the vague topology on the total variation ball $B_{\mathcal{M}}(f_1(x_0))$, we can extract a subsequence $\lambda_{n_k}^*$ which converges vaguely to some measure $\lambda_\infty^* \in B_{\mathcal{M}}(f_1(x_0)) \cap \mathcal{M}^+(\Lambda)$. The function $\widetilde{g}$ lives in $C_0(\Lambda)$ and consequently we obtain $\lim_{k\to\infty} d_{n_k} = \langle \widetilde{g}, \lambda_\infty^* \rangle$. From above we know that the sequence of optimal values $p_n = d_n$ is decreasing and bounded from below. This shows that $\lim_{n\to\infty} p_n = \lim_{n\to\infty} d_n = \langle \widetilde{g}, \lambda_\infty^* \rangle$ holds true, as claimed. Assumption (3.9) warrants that $\widetilde{V}_l \in C_0(\Lambda)$ for any $l \in \{2, 3, ...\}$ and therefore

$$\langle \widetilde{V}_l, \lambda_\infty^* \rangle = \lim_{k\to\infty} \langle \widetilde{V}_l, \lambda_{n_k}^* \rangle \leq V(T, x_0).$$

Clearly, we have $\langle \widetilde{V}_1, \lambda_\infty^* \rangle = \|\lambda_\infty^*\| \leq f_1(x_0) = V(T, x_0)_1$ which finally implies that λ_∞^* is indeed $\mathrm{D}_n^{\mathrm{c}\prime}$-admissible.

4. *This assertion may be obtained as a consequence of some profound results from optimization theory. Nevertheless, a short argument is provided for didactic purposes. The reader is kindly asked to pardon this mathematical detour.*

We will show that it is always possible to choose a dual optimizer which is supported on $n+1$ points of the set Λ. Due to the growth conditions (3.9,3.10) and the invariance of the payoff f_1 with respect to the pricing semi-group, the mapping $\phi := (\widetilde{g}, \widetilde{V}) : \Lambda \to \mathbb{R}^{n+1}$ extends continuously to the Alexandorff compactification Λ_∞ of the set Λ, cf. [AL, Theorem 2.72]. Consequently, there exist compact sets $D \subset \mathbb{R}, E \subset \mathbb{R}^{n-1}$ such that

$$\phi(\Lambda_\infty) = D \times \{1\} \times E.$$

Obviously, the vector $\frac{1}{\|\lambda_n^*\|}\langle \phi, \lambda_n^* \rangle$ is contained in the closure of the convex hull of the set $\phi(\Lambda_\infty)$. Basic calculus yields that $\mathrm{co}(\phi(\Lambda_\infty))$ is a compact subset of $\mathbb{R}^{n+1}$. By virtue of Caratheodory's theorem, cf. [RO2, Theorem 17.1], there exist $(\vartheta_i, x_i)_{i=1,...,n+1} \subset \Lambda$ and $\alpha \in \mathbb{R}_+^{n+1}$ such that

$$\langle \phi, \lambda_n^* \rangle = \|\lambda_n^*\| \sum_{i=1}^{n+1} \alpha_i \phi(\vartheta_i, x_i) = \langle \phi, \nu \rangle \tag{3.14}$$

where $\nu := \|\lambda_n^*\| \sum_{i=1}^{n+1} \alpha_i \delta_{(\vartheta_i, x_i)}$. Clearly, the latter equation implicitly warrants that $\sum_{i=1}^{n+1} \alpha_i = 1$. Please, give a brief thought to the fact that we indeed only require at most $n+1$ summands in the convex combination from (3.14) and that $(\vartheta_i, x_i)_{i=1,\dots,n+1}$ can be chosen from the set Λ, i.e. the measure ν assigns no mass to ∞. Equation (3.14) can be equivalently restated as

$$\begin{aligned}\langle \widetilde{g}, \nu\rangle &= \langle \widetilde{g}, \lambda_n^*\rangle \\ \langle \widetilde{V}, \nu\rangle &= \langle \widetilde{V}, \lambda_n^*\rangle\end{aligned}$$

and therefore the measure ν is $\mathrm{D}_n^{\mathrm{c}\,\prime}$-optimal.

□

From Proposition 3.4 it is apparent that the existence of a $\mathrm{P}_n^{\mathrm{c}}$-optimal vector $a^{*,n}$ does not depend on Assumption (3.9). The latter was crucial in order to establish the existence of a dual optimizer. Combining Proposition 3.4 and Lemma 3.8 we obtain the following consistency result. Again, we write $\langle\,\cdot,\lambda\rangle := \int_{[0,T]\times\mathbb{R}} \cdot\, \mathrm{d}\lambda$ for any $\lambda \in \mathcal{M}^+([0,T]\times\mathbb{R})$.

3.15 Lemma: *Let $f_1(x) := e^{-\frac{2r}{\sigma^2}x} + e^x$ and suppose that $f_2, f_3, \dots : \mathbb{R} \to \mathbb{R}_+$ are continuous functions such that $f_1, f_2, \dots$ are linearly independent. Assume that $v_{\mathrm{eu},f_k}(T+\varepsilon, x_0) < \infty$ for some $\varepsilon > 0$ and that $f_2, f_3, \dots$ satisfy the growth condition (3.9). Furthermore, let $g : \mathbb{R} \to \mathbb{R}_+$ denote a continuous American payoff which satisfies (3.10). Clearly, for any $n \in \mathbb{N}$ the linear program P_n^c is equivalent to P_n with $\mathrm{d}\mu_k(y) := \mathrm{N}\left(x_0 + \hat{r}T, \sigma^2 T, y\right) f_k(y)\mathrm{d}y$ and $\widetilde{U}_n := \{\sum_{k=1}^n a_k \mu_k \mid a \in \mathbb{R}_+^n\}$. Suppose that there exist a sequence $\xi_n \in \widetilde{U}_n$ which suffices the assumptions from the third assertion of Proposition 3.4. Then:*

1. *For any $n \in \mathbb{N}$ the optimal value p_n in Program P_n is attained by some measure $\mu_n^* = \sum_{k=1}^n a^{*,n} \mu_k$ and the optimal value d_n in D_n^c is attained by some admissible measure λ_n^*. We have strong duality $p_n = d_n$ and the complementary slackness equations* (3.11) *hold true.*

2. *There exists a subsequence n_k such that $\mu_{n_k}^*$ converges weakly to some P_∞-optimal measure ν_∞ and $\lambda_{n_k}^*$ converges vaguely to some measure λ_∞^* which is D_n^c-admissible for any $n \in \mathbb{N}$.*

3. *The optimal values satisfy*

$$p_\infty = \lim_{n\to\infty} p_n = \lim_{n\to\infty} d_n = \langle g, \lambda_\infty^*\rangle$$

and the limiting complementary slackness property

$$v_{\mathrm{eu},\nu_\infty} = g \qquad \lambda_\infty^*\text{-a.e. on } (0,T)\times\mathbb{R} \tag{3.16}$$

holds true.

Proof.

1. This follows directly from the first and second assertion of Lemma 3.8.

2. The third assertion of Proposition 3.4 warrants that there exists some subsequence $\mu^*_{n_l}$ which converges weakly to some P_∞-optimal measure ν_∞ with optimal value $p_\infty = \lim_{n\to\infty} p_n$. By further thinning out the sequence n_l, we can obtain a subsequence n_k such that the dual optimizers $\lambda^*_{n_k}$ converge vaguely to some measure λ^*_∞ which is D^c_n-admissible for any $n \in \mathbb{N}$. The argument is literally the same as in the proof of Lemma 3.8, Assertion 3. Furthermore, we have $\lim_{n\to\infty} d_n = \langle g, \lambda^*_\infty\rangle$.

3. From above we already know that $p_\infty = \lim_{n\to\infty} p_n = \lim_{n\to\infty} d_n = \langle g, \lambda^*_\infty\rangle$. Let us verify Equation (3.16). The admissibility of the measures $\mu^*_{n_k}$ and ν_∞ in the respective primal programs warrant that the quantities $v_{\mathrm{eu},\mu^*_{n_k}} - g$ and $v_{\mathrm{eu},\nu_\infty} - g$ are non-negative on the set $(0,T)\times\mathbb{R}$. The subsequence $\mu^*_{n_k}$ converges weakly to ν_∞ and therefore we can apply Fatou's lemma in order to obtain
$$0 \le \int_{(0,T)\times\mathbb{R}} v_{\mathrm{eu},\nu_\infty} - g \,\mathrm{d}\lambda^*_\infty \le \liminf_{k\to\infty} \int_{(0,T)\times\mathbb{R}} v_{\mathrm{eu},\mu^*_{n_k}} - g \,\mathrm{d}\lambda^*_\infty.$$
Due to the fact that the measure λ^*_∞ is D^c_n-admissible for any $n \in \mathbb{N}$, we find that
$$\begin{aligned}\int_{(0,T)\times\mathbb{R}} v_{\mathrm{eu},\mu^*_{n_k}} - g \,\mathrm{d}\lambda^*_\infty &\le \int_{[0,T]\times\mathbb{R}} v_{\mathrm{eu},\mu^*_{n_k}} - g \,\mathrm{d}\lambda^*_\infty \\ &= \sum_{l=1}^{n_k} a_l^{*,n_k} \langle v_{\mathrm{eu},f_l}, \lambda^*_\infty\rangle - \langle g, \lambda^*_\infty\rangle \\ &\le \sum_{l=1}^{n_k} a_l^{*,n_k} v_{\mathrm{eu},f_l}(T, x_0) - \langle g, \lambda^*_\infty\rangle\end{aligned}$$
holds true. Clearly, the vector a^{*,n_k} is optimal in Program P^c_n which yields that $\sum_{l=1}^{n_k} a_l^{*,n_k} v_{\mathrm{eu},f_l}(T,x_0) = p_{n_k}$. Finally, we obtain
$$\int_{(0,T)\times\mathbb{R}} v_{\mathrm{eu},\nu_\infty} - g \,\mathrm{d}\lambda^*_\infty = \liminf_{k\to\infty} p_{n_k} - \langle g, \lambda^*_\infty\rangle = 0$$
which shows that (3.16) is indeed valid. □

For any vector $a \in \mathbb{R}^n_+$ which is admissible in P^c_n define $f[a] := \sum_{k=1}^n a_k f_k$. The latter payoff is superreplicating in the sense of Definition 2.1 and in virtue of Proposition 2.3, we have
$$v_{\mathrm{am},g}(\vartheta, x) \le \sum_{k=1}^n a_k \, v_{\mathrm{eu},f_k}(\vartheta, x) = v_{\mathrm{eu},f[a]}(\vartheta, x) \tag{3.17}$$
for any $(\vartheta, x) \in [0,T]\times\mathbb{R}$. In other words, the value function associated to the European payoff $f[a]$ always provides us with a global bound for the American value function associated to g. In our numerical experiments we applied the cutting plane procedure described in [LW] and [IW] in order to approximate an optimizer of program P^c_n. Let us outline a basic version of the algorithm:

3.18 Algorithm:

Step 1: Choose some finite set $\Gamma_1 \subset [0,T] \times \mathbb{R}$ of initial constraints, fix some maximal number of iterations $m_{\max} \in \mathbb{N}$ and put $m := 1$.

Step 2: Calculate the solution $a^{(m)} \in \mathbb{R}_+^n$ of the following finite dimensional linear program:

$$\begin{array}{lll} \text{minimize} & V(T,x_0)^\top a & (\mathrm{P}_{\Gamma_m}) \\ \text{subject to} & a \in \mathbb{R}_+^n, & \\ & V(\vartheta,x)^\top a \geq g(x) \text{ for any } (\vartheta,x) \in \Gamma_m & \end{array}$$

Step 3: Calculate the point $(\vartheta^{(m)}, x^{(m)}) \in [0,T] \times \mathbb{R}$ where the superreplication constraint is most severely violated, i.e.

$$(\vartheta^{(m)}, x^{(m)}) := \underset{(\vartheta,x)\in[0,T]\times\mathbb{R}}{\operatorname{argmin}} \left\{ V(\vartheta,x)^\top a^{(m)} - g(x) \right\}.$$

Step 4: Put $\Gamma_{m+1} := \Gamma_m \cup \{(\vartheta^{(m)}, x^{(m)})\}$.

Step 5: If the maximum number of iterations is reached, i.e. $m = m_{\max}$, or some other prespecified break criterion is satisfied, output

$$\tilde{a} := a^{(m)}$$

as approximate solution of $\mathrm{P}_n^{\mathrm{c}}$ and terminate the algorithm. Otherwise, increase the iteration counter $m := m+1$ and return to the second step.

Here we use the notation $V(\vartheta,x) := (v_{\mathrm{eu},f_1}(\vartheta,x), ..., v_{\mathrm{eu},f_n}(\vartheta,x)) \in \mathbb{R}^n$ from above. Given the weight vector $\tilde{a}$, we obtain the approximate CDEO $\tilde{f} := f[\tilde{a}] = \sum_{k=1}^n a_k^{(m)} f_k$. From [LW, Theorem 5.1] we know that a subsequence of the P_{Γ_m}-optimal elements $a^{(m)}$ converges to a solution of program $\mathrm{P}_n^{\mathrm{c}}$ as $m \to \infty$. Irrespective of representability issues, the practitioner might therefore consider $v_{\mathrm{eu},\tilde{f}} = v_{\mathrm{eu},f[a^{(m)}]}$ as a global upper bound for the American value function $v_{\mathrm{am},g}$ within the limits of numerical accuracy if m is sufficiently large, see also Section 3.3 on this behalf.

In our numerical experiments it turned out to be convenient to generate the set Γ_1 from *Step 1* of Algorithm 3.18 randomly by drawing independent samples from some probability law on $[0,T] \times \mathbb{R}$. Moreover, in [IW, Algorithm 2.2] a slight modification of the cutting plane procedure is suggested in order to enhance the performance of the algorithm. The key idea is to drop in each iteration the inactive constraints from the set Γ_m which possibly reduces the overall computation time as the dimension of the linear program $\mathrm{P}_{\Gamma_{m+1}}$ might decrease. We observed that sometimes it is even more efficient to allow for a short burn-in period before dropping the inactive constraints. In other words, Algorithm 3.18 may be modified as follows: Let $m_{\text{burn-in}} \in \mathbb{N}_0$ denote some prespecified number of burn-in iterations. Replace *Step 4* of Algorithm 3.18 by

Step 4': Let $\lambda^m \in \mathbb{R}_+^{\Gamma_m}$ denote a solution of the Lagrange dual associated to the linear program P_{Γ_m}, i.e. λ^m is an extremal element of the maximization problem

$$\begin{aligned} &\text{maximize} && \sum_{(\vartheta,x)\in\Gamma_m} g(x)\lambda(\vartheta,x) && (\mathrm{D}_{\Gamma_m}) \\ &\text{subject to} && \lambda \in \mathbb{R}_+^{\Gamma_m}, \\ &&& \sum_{(\vartheta,x)\in\Gamma_m} V(\vartheta,x)\lambda(\vartheta,x) \leq V(T,x_0). \end{aligned}$$

If the number of burn-in iterations is exceeded, i.e. $m > m_{\text{burn-in}}$, put

$$\Gamma_{m+1} := \{(\vartheta,x) \in \Gamma_m \mid \lambda^m(\vartheta,x) > 0\} \cup \{(\vartheta^{(m)}, x^{(m)})\}.$$

Otherwise, put $\Gamma_{m+1} := \Gamma_m \cup \{(\vartheta^{(m)}, x^{(m)})\}$.

As discussed above, numerically solving the general capacity problem GCAP involves two crucial steps: First, a suitable discretization of GCAP needs to be chosen in order to approximate the latter optimization task by a sequence of semi-infinite linear programs. If the CDEO corresponds to a weakly approximable solution of Program P_∞ in the sense of Assertion 3 from Proposition 3.4, we can select a subsequence of primal optimizers which converges weakly to a measure-type CDEO ν_∞. To put differently, increasing the number of basis elements which are taken into account by the semi-infinite program, ultimately leads to better approximations of ν_∞.
Secondly, we execute the iterative cutting plane procedure, as outlined in Algorithm 3.18, in order to solve the semi-infinite approximation to GCAP within the limits of numerical accuracy. The practical implementation of this procedure leaves plenty of room for problem specific adjustments. For example, the performance of Algorithm 3.18 may be increased by choosing a discretization which makes use of the geometric structure of the American payoff or reflects some prior information about the potential CDEO. Below, we will present a detailed analysis of the American put within the scope of the Black-Scholes market which ultimately leads to a fast and precise pricing method, cf. Sections 3.2 and 3.3.

3.2 A qualitative study: Representability of American put options

In [JM2] it was conjectured that the American put is not representable by any European claim. In this paragraph we want to provide some numerical results which raise the hope that the American put $g(x) = (e^K - e^x)_+$ is representable in the Black-Scholes market (1.3) after all.

Theorem 2.51 warrants that the CDEO associated to the put exists as a measure which is concentrated on the interval $(-\infty, K]$. For the study contained in this section, we consider a discretization of GCAP which does not rely on any further prior information concerning the geometric structure of the CDEO. To be more specific, the density functions $f_2, ..., f_n$ occurring in Program $\mathrm{P}_n^{\mathrm{c}}$ were randomly generated in the following manner: We fix some small constant $\varepsilon > 0$ and put $x_1 = -\infty$ and $x_n := K - \varepsilon$. Afterwards we randomly generate $x_2 < ... < x_{n-1}$ by independently drawing $n-2$ numbers from the probability distribution ν with Lebesgue density

$$\frac{\mathrm{d}\nu}{\mathrm{d}x} = e^{x-K+\varepsilon}\mathbb{1}_{(-\infty,K-\varepsilon)}(x). \tag{3.19}$$

We define $f_1(x) := e^{-\frac{2r}{\sigma^2}x} + e^x$ and for any $k \in \{2, ..., n\}$ we choose f_k to be the bump function as depicted in Figure 3.1, i.e.

$$f_k(x) := \begin{cases} \frac{1}{\varepsilon}(x - x_{k-1} + \varepsilon) & \text{if } x \in (x_{k-1} - \varepsilon, x_{k-1}), \\ 1 & \text{if } x \in [x_{k-1}, x_k], \\ \frac{1}{\varepsilon}(x_k + \varepsilon - x) & \text{if } x \in (x_k, x_k + \varepsilon), \\ 0 & \text{otherwise.} \end{cases} \tag{3.20}$$

Here we follow the convention that $(-\infty, -\infty) := \emptyset$. Lemma 3.8 warrants that the optimal values in $\mathrm{P}_n^{\mathrm{c}}$ and $\mathrm{D}_n^{\mathrm{c}}$ are attained and that strong duality as well as the complementary slackness equations from (3.11) hold true.

Algorithm 3.18 is applied in order to generate an approximate CDEO $\tilde{f}$. Our computations are performed with Matlab R2014a on a standard home computer with an Intel Core i3-3240 CPU. The linear program P_{Γ_m} is solved using the Matlab routine `linprog`. For the non-linear minimization task from *Step 3* we use the solver `fmincon` from the Matlab optimization toolbox. The implementation of Algorithm 3.18 is very simple and requires only a few lines of source code.

Let us consider the market (1.3) with the parameters $T = 0.5$, $r = 0.06$, $\sigma = 0.4$ and the put payoff $g(x) = (e^K - e^x)_+$ with log-strike price $K = \log(100)$. In favor of a clearly arranged visualization, we choose a non-logarithmic price coordinate $s = e^x$ for the presentation of our numerical data. The reader may adequately reformulate the theorems of Section 2.2 for non-logarithmic stock prices. Figure 3.2 depicts the graph of the

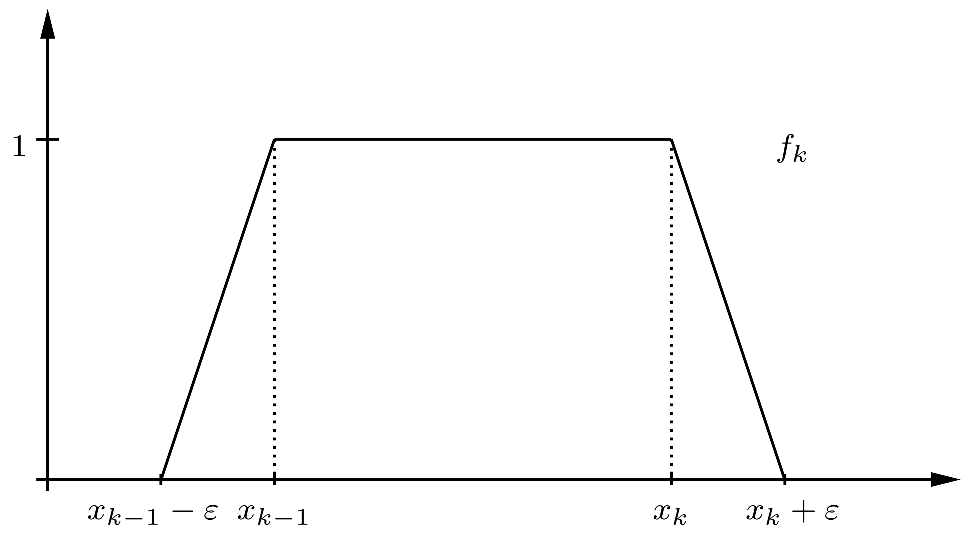

Figure 3.1: The bump function f_k as defined in (3.20).

European value function associated to the approximate CDEO $\tilde{f}$ obtained by Algorithm 3.18 using the optimization point $(T, x_0) = (0.5, \log(100) + 0.1)$. The s-axis represents the non-logarithmic stock price of the underlying and the ϑ-axis indicates the maturity of the option. The transparent surface in Figure 3.2 corresponds to the American put payoff.

Now we present some numerical results which show that from a qualitative point of view the requirements of Theorem 2.52 seem to be satisfied within the limits of numerical accuracy. Taking Figure 3.2 into account, we see that there is no explosion at the $\vartheta = T$ border and that $\liminf_{\vartheta \to 0} v_{\mathrm{eu},\tilde{f}}(\vartheta, e^K) = 0$. To put differently, our data indicates that the first and fourth assumption of the aforementioned theorem are satisfied. The white curve $\vartheta \mapsto \widehat{s}(\vartheta)$ depicted in the Figure 3.3 corresponds to the unique minima of the mappings $(0, e^K] \ni s \mapsto v_{\mathrm{eu},\tilde{f}}(\vartheta, \log(s)) - g(\log(s))$. Qualitatively, the requirement $\liminf_{\vartheta \to 0} \widehat{s}(\vartheta) = 100 = e^K$ from Assumption 2 indeed seems to hold true. Moreover, Figure 3.4 depicts the mapping $(0, T] \ni \vartheta \mapsto H(\vartheta, \log \widehat{s}(\vartheta))$ where H is defined as in (2.53). We see that $H(\vartheta, \log \widehat{s}(\vartheta)) \approx 75 = \frac{2r}{\sigma^2} e^K > 0$ for any $\vartheta \in (0, T]$ and therefore the third requirement of Theorem 2.52 seems to be satisfied as well. From an engineer's perspective, these findings might be considered as an indication that - notwithstanding the studies of [JM2] - the American put is representable within the scope of the Black-Scholes model.

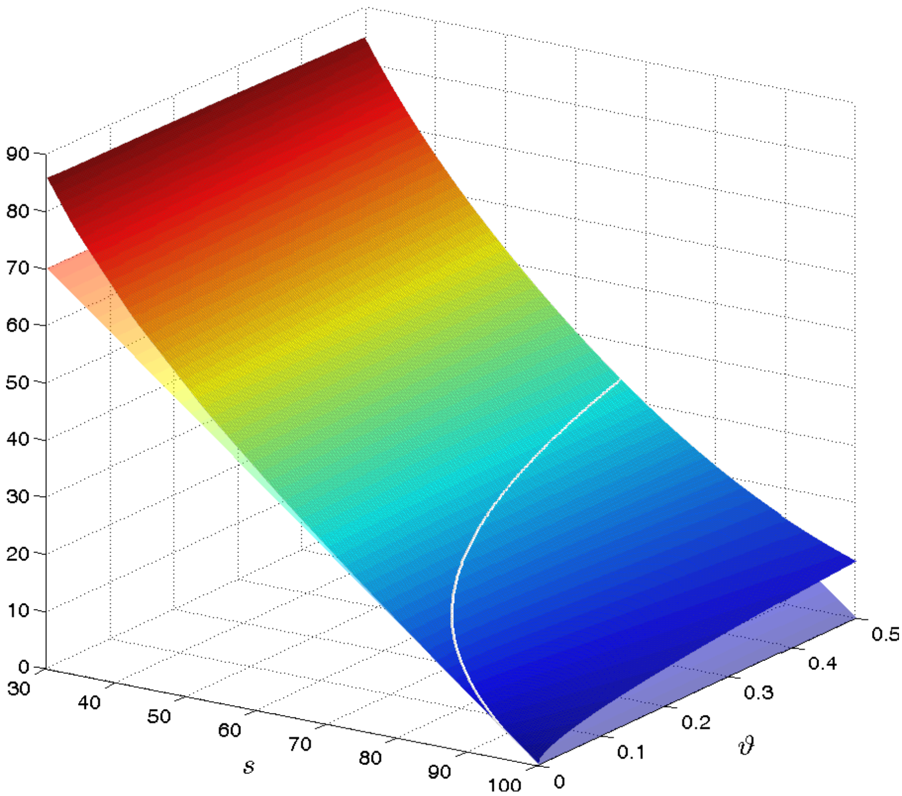

Figure 3.2: The price surface of the approximate CDEO associated to the Am. put.

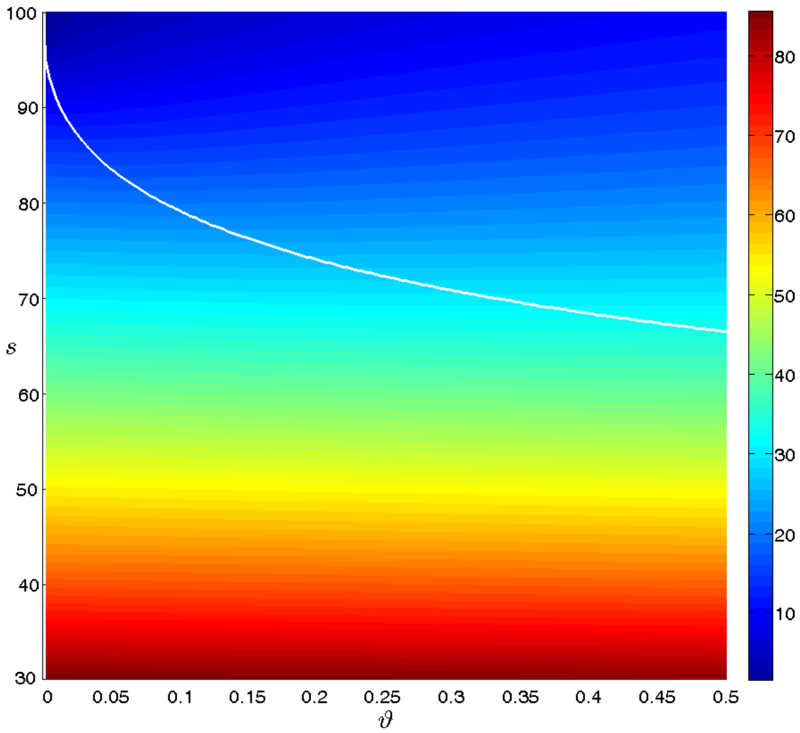

Figure 3.3: The curve $\widehat{s}$ associated to the approximate CDEO of the Am. put.

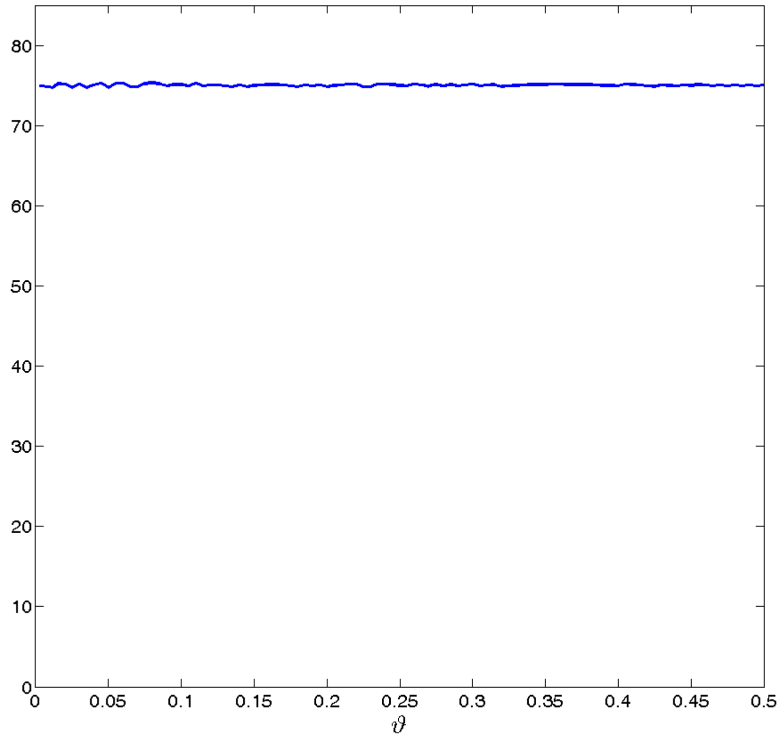

Figure 3.4: The mapping $H(\cdot, \log \widehat{s}(\cdot))$ associated to the approximate CDEO of the American put.

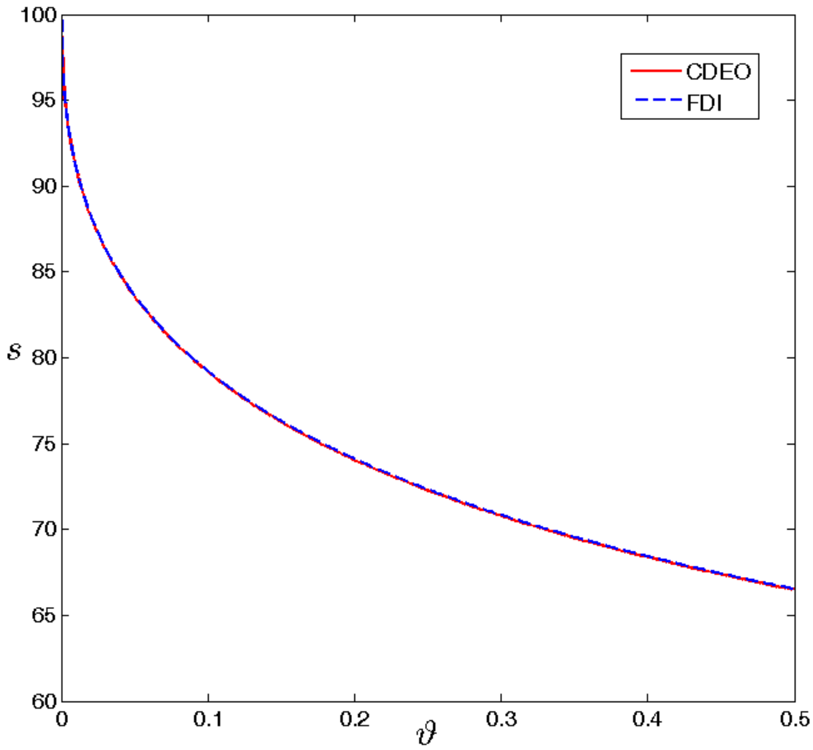

Figure 3.5: Comparison between the CDEO minima curve $\widehat{s}$ and a FDI approximation to the early exercise boundary associated to the American put.

Let us remark that the numerical data obtained from our simulation indeed complies with the theoretical implications of the put being representable. For example, for any $x \in \operatorname{cl}\pi(\widetilde{C}_{(T,x_0)}) \approx (\log(66.5), \infty)$ we have that $g(x) = \inf_{\vartheta \in [0,T]} v_{\mathrm{eu},\tilde{f}}(\vartheta, x)$ holds true within the working accuracy of Algorithm 3.18. Compare this observation to Assertion (d) of Theorem 2.52. Moreover, we generated an approximation to the early exercise boundary using a Crank-Nicolson finite difference scheme on a very fine grid, see Figure 3.5. As anticipated by Assertion (e) of Theorem 2.52, the latter approximation matched the curve $\widehat{s}$ obtained by the CDEO method within the working accuracy of the algorithm. For further comparisons between the CDEO approach and different pricing methods from the literature, we refer the reader to Section 3.3.

Summing up, for the American put in the Black-Scholes market all qualitative requirements of Theorem 2.52 are satisfied from a numerical viewpoint. Furthermore, we remark that a detailed analysis of the European price surface obtained in our numerical experiment indicates that the approximate CDEO is not contained in the class of European payoffs which was proposed by [JM2]. In opposition to [JM2], we therefore have hope that the American put is representable by a European claim after all. Moreover, Figure 3.2 indicates that within the scope of the Black-Scholes model the potential CDEO associated to the American put might be well approximated by a decreasing, continuous, convex mapping which vanishes on the interval (e^K, ∞). This prior information can be used beneficially in order to increase the performance of the pricing procedure, cf. Section 3.3 below.

3.3 Speed and precision of CDEO price approximations

Again, we consider the Black-Scholes market (1.3) and some American payoff function $g : \mathbb{R} \to \mathbb{R}_+$. We follow the convention from above that x and $s = e^x$ denote logarithmic and non-logarithmic stock prices, respectively.

Independent of the question whether g is representable, the method outlined in Section 3.1 can be used to compute upper bounds of the American option price. More precisely, Algorithm 3.18 can be applied in order to generate an approximate solution $\widetilde{a}$ of Program $\mathrm{P}_n^{\mathrm{c}}$. The associated European payoff $\widetilde{f} := \sum_{k=1}^{n} \widetilde{a}_k f_k$ superreplicates g up to time T within the limits of numerical accuracy. Hence, within this range of precision, the associated European value function $v_{\mathrm{eu},\tilde{f}}$ constitutes a global upper bound of the American value function associated to g on the set $[0, T] \times \mathbb{R}$, see Equation (3.17). Moreover, in case that g is representable relative to (T, x_0) by some measure μ^*, Lemma 3.2 warrants that the latter coincides with the cheapest dominating European option. Furthermore, if the discretization under consideration is sufficiently rich to approximate μ^* in the sense of Assertion 2 from Lemma 3.7, the latter indicates that $v_{\mathrm{eu},\tilde{f}}$ may indeed constitute a reasonable approximation to $v_{\mathrm{am},g}$ on the connected component $C_{(T,x_0)}$.

In this section we will apply the CDEO method to American put options in the Black-Scholes setting for different model configurations. The performance and accuracy of the pricing procedure will be assessed by comparing the latter with classic approaches and high performance methods from the literature. All computations were performed with Matlab R2014a on a standard home computer using one core of an Intel Core i3-3240 processor. In particular, no parallel processing techniques were used in order to increase the computation speed. The linear program P_{Γ_m} was solved using the Matlab routine `linprog` with the `simplex` algorithm. The non-linear minimization task from *Step 3* of Algorithm 3.18 was solved using `fmincon` with the `sqp` algorithm from the Matlab optimization toolbox. Both solvers were configured to work with an objective function tolerance of 10^{-8}. Moreover, the `sqp` algorithm was configured to work with a constraint and step tolerance of 10^{-8}. Our numerical experiments demonstrate that the price approximations of the CDEO method easily achieve accuracies in the order of the computational precision of the subroutines `fmincon` and `linprog` ($= 10^{-8}$). Consequently, we require high precision reference values in order to assess the performance of the CDEO algorithm. Aside from the price approximations of [ALO], we could not find other sufficiently accurate numerical results in the literature. For this reason, the classical binomial tree method and a finite difference scheme were implemented in order to generate more benchmark values.

Based on a preceding qualitative analysis of the potential CDEO associated to the American put for different parameter sets, we choose in Algorithm 3.18 a discretization which is advantageous for representing functions of the type $x \mapsto c(e^x)$ where $c : \mathbb{R}_+ \to \mathbb{R}_+$ denotes a decreasing, convex mapping with $c(s) = 0$ for any $s > e^K$. On this behalf, the reader may want to re-examine the price surface depicted in Figure 3.2 and recall the short remark from the end of Section 3.2. By virtue of [NP, Theorem 1.6.3], there exists a uniquely determined measure μ concentrated on $[0, e^K]$ such that

$$c(e^x) = \int_{[0,e^K]} (y - e^x)_+ \mathrm{d}\mu(y)$$

holds true for any $x \in \mathbb{R}$. The latter representation formula motivates choosing $f_k(x) := (a_k - e^x)_+$ in Program $\mathrm{P}_n^{\mathrm{c}}$ where $0 < a_1 < ... < a_{n-1} < a_n := e^K$ are generated by independently drawing $n-1$ numbers from the uniform distribution on $(0, e^K)$. Thanks to the Black-Scholes formula, any European price $v_{\mathrm{eu},f_k}(\vartheta, x)$ appearing in Algorithm 3.18 can be easily computed. In any of the test cases below, this tailor-made discretization which reflects the properties of the potential CDEO indeed outperforms the all-purpose discretization consisting of bump functions from Section 3.2.

Test 1: We consider the Black-Scholes market (1.3) with the parameters $r = 0.06$ and $\sigma = 0.4$ and assume that the risky asset does not pay any dividends. The CDEO method, binomial trees and a finite difference scheme are applied in order to approximate the value of an American put option with strike 100 at spot price $s = 100$ and maturity $\vartheta = 0.5$. Following the description from above, we randomly generate $n := 1500$ basis functions of the type $f_k(x) := (a_k - e^x)_+$. The set Γ_1 from *Step 1* of Algorithm 3.18 is created by drawing 225 independent samples from the distribution $\mathcal{L}(\mathcal{U}_{[0,0.5]}) \otimes \mathcal{L}(\log \mathcal{U}_{[50,100]})$. Recall that $\mathcal{U}_{[a,b]}$ denotes a random variable which is uniformly distributed on the interval $[a, b]$. Moreover, we choose a to terminate the procedure after $m_{\max} := 100$ iterations. In order to keep the linear sub-problems P_{Γ_m} small, we start to drop the inactive constraints after one burn-in iteration, i.e. we choose $m_{\text{burn-in}} := 1$ in *Step 4'* of Algorithm 3.18. Table 3.1 shows how the approximation to the American option price, i.e. the optimal value of the linear sub-problem P_{Γ_m}, evolves as the number of iterations m increases. The column *total time* indicates the total amount of time (measured in seconds) that has passed since the algorithm was started.

Table 3.2 contains the American option prices which were obtained from differently sized binomial tree approximations. We refer the reader to [CRR] and [SD, Section 1.4] for a detailed exposition of this classic approach towards American option pricing. The column *time steps* indicates how many time steps were used in the discrete binomial tree approximation to the Black-Scholes model. The column *Am. value* contains value of the American put in the corresponding binomial tree model which serves as approximation to the true option price. Moreover, the column *diff* indicates the absolute difference between the values contained in the column *Am. value* and the $m = 100$ CDEO price approximation (= 9.9451361327103609). The time (measured in seconds) that was needed to perform the binomial tree algorithm with a specific number of time steps is indicated in the rightmost column of the table.

iteration m	Am. value	total time (s)
1	9.9408517316167142	1.91
2	9.9421725315439708	2.03
3	9.9429005244135009	2.16
4	9.9441737005523354	2.27
5	9.9446608846442341	2.41
10	9.9447909897334164	2.98
15	9.9450419755448927	3.59
20	9.9450918945713589	4.32
25	9.9451195188610768	5.03
30	9.9451326596020539	5.67
35	9.9451329908739599	6.33
40	9.9451331298143426	7.07
45	9.9451332558993855	7.73
50	9.9451345475142148	8.38
55	9.9451349276089367	9.07
60	9.9451357041082247	9.76
65	9.9451357150669732	10.45
70	9.9451358487366477	11.22
75	9.9451360155812747	12.03
80	9.9451360335244097	12.81
85	9.9451360717728186	13.50
90	9.9451361425532507	14.27
95	9.9451360704801139	14.98
100	9.9451361327103609	15.72

Table 3.1: CDEO method: Am. put, strike 100, spot $s = 100$, maturity $\vartheta = 0.5$. BS parameters: $r = 0.06, \sigma = 0.4$, no dividends.

time steps	Am. value	diff	time (s)
5000	9.944912910948	2.23e-04	0.22
10000	9.945024842240	1.11e-04	0.87
15000	9.945062144882	7.40e-05	2.10
20000	9.945080667687	5.55e-05	3.88
25000	9.945091868237	4.43e-05	6.17
30000	9.945099297771	3.68e-05	8.97
35000	9.945104621245	3.15e-05	12.30
40000	9.945108596040	2.75e-05	16.26
45000	9.945111672723	2.45e-05	20.78
50000	9.945114148085	2.20e-05	26.45
250000	9.945131919148	4.21e-06	1282.28
500000	9.945134147094	1.99e-06	5899.42
750000	9.945134878000	1.25e-06	13782.69
1000000	9.945135254429	8.78e-07	24804.74

Table 3.2: Binomial tree method: Am. put, strike 100, spot $s = 100$, maturity $\vartheta = 0.5$. BS parameters: $r = 0.06, \sigma = 0.4$, no dividends.

Table 3.3 contains price approximations which were obtained by applying a simple finite difference method. The linear complementary problem (1.2) associated to the valuation of the American put was discretized using the Crank-Nicolson scheme, cf. [SD, Subsection 4.6.1]. The resulting finite dimensional linear complementary problems were solved by the Brennan-Schwartz algorithm in each time step, cf. [TV]. In order to reduce the memory consumption of the finite difference approach, we only stored the price vector associated to the current iteration. The column *Am. value* displays the approximate American put price generated by the finite difference method for differently sized grids. Here we denote by N and $M = 25N$ the number of grid points in the spatial and temporal dimension, respectively. As above, the column *diff* contains the absolute difference between the values of the column *Am. value* and the $m = 100$ CDEO price approximation. Moreover, the column *time* indicates how many seconds were needed in order to execute the finite difference method on the corresponding grid.

gridsize	M	N	Am. value	diff	time (s)
6.2500e+06	12500	500	9.9445174532114962	6.19e-04	0.50
2.5000e+07	25000	1000	9.9451998227344180	6.37e-05	1.42
1.0000e+08	50000	2000	9.9451474879326742	1.14e-05	4.58
2.2500e+08	75000	3000	9.9451392454009326	3.11e-06	9.61
4.0000e+08	100000	4000	9.9451368331704000	7.00e-07	16.25
6.2500e+08	125000	5000	9.9451359283724994	2.04e-07	25.76
9.0000e+08	150000	6000	9.9451355515935322	5.81e-07	35.48
1.2250e+09	175000	7000	9.9451353928201858	7.40e-07	47.33
1.6000e+09	200000	8000	9.9451353324134786	8.00e-07	60.81
2.0250e+09	25000	9000	9.9451353219062959	8.11e-07	77.26
2.5000e+09	250000	10000	9.9451353352243199	7.97e-07	94.23
3.0250e+09	275000	11000	9.9451353625823185	7.70e-07	113.66
3.6000e+09	300000	12000	9.9451353946756598	7.38e-07	134.92
6.4000e+09	400000	16000	9.9451355310030074	6.02e-07	233.98
1.2100e+10	550000	22000	9.9451356935361730	4.39e-07	451.56
1.9600e+10	700000	28000	9.9451358085134220	3.24e-07	757.28
2.8900e+10	850000	34000	9.9451358898906204	2.43e-07	1124.13
4.0000e+10	1000000	40000	9.9451359551683645	1.78e-07	1554.80
5.2900e+10	1150000	46000	9.9451359961004719	1.37e-07	2079.19
6.7600e+10	1300000	52000	9.9451360333156096	9.94e-08	2627.33
8.4100e+10	1450000	58000	9.9451360741420523	5.86e-08	3277.86
1.0240e+11	1600000	64000	9.9451360960087598	3.67e-08	3984.91
1.2250e+11	1750000	70000	9.9451361160198779	1.67e-08	4760.19

Table 3.3: FDI method: Am. put, strike 100, spot $s = 100$, maturity $\vartheta = 0.5$. BS parameters: $r = 0.06, \sigma = 0.4$, no dividends.

Test 2: We compare the CDEO pricing scheme to the high-precision approach put forward by [ALO]. The authors of the latter article consider a Black-Scholes market where the risky asset pays dividends at a continuous rate of q, i.e. we have an exponential type riskless bond $B_t = \exp(rt)$ and the price of the risky asset evolves according to the stochastic differential equation $\mathrm{d}S_t = (r-q)S_t\mathrm{d}t + \sigma S_t \mathrm{d}W_t$.

2.a: First, we consider the market parameters $r = q = 0.05, \sigma = 0.25$. The task is to calculate the American premium of an American put option with strike 100 at spot price $s = 100$ and maturity $\vartheta = 1$. For this purpose, we initialize Algorithm 3.18 with $n := 1800$ randomly generated basis functions of the type $f_k(x) := (a_k - e^x)_+$. The set Γ_1 is created by independently drawing 400 samples from the distribution $\mathcal{L}(\mathcal{U}_{[0,1]}) \otimes \mathcal{L}(\log \mathcal{U}_{[50,100]})$. Again, we decide to drop the inactive constraints after one burn-in iteration in order to reduce the size of the linear sub-problems P_{Γ_m}. Table 3.4 shows how the CDEO approximation to the American option value and the American premium evolves as the number of iterations m increases. The column *total time* indicates the total amount of seconds that has passed since the algorithm was started. We compare the CDEO approximation to the American premium against the $(l, m, n) = (65, 8, 32)$ value from [ALO, Table 2] which is claimed to be accurate to about 12 digits. The absolute difference between the values contained in the column *Am. premium* and the aforementioned reference premium is displayed in the column *diff*. After 75 iterations the order of this discrepancy matches the preset computational accuracy of the solvers `linprog` and `fmincon`.

iteration m	Am. value	Am. premium	diff	total time (s)
1	9.558525193173	0.096032597006	1.09e-02	1.10
5	9.564762795109	0.102270198941	4.68e-03	1.81
10	9.564781106704	0.102288510537	4.66e-03	2.89
15	9.566509074981	0.104016478814	2.94e-03	3.82
20	9.569151771488	0.106659175320	2.94e-04	4.63
25	9.569404749893	0.106912153726	4.05e-05	5.54
30	9.569424462819	0.106931866652	2.08e-05	6.50
35	9.569431698379	0.106939102212	1.36e-05	7.61
40	9.569439598899	0.106947002732	5.70e-06	8.92
45	9.569441847972	0.106949251805	3.45e-06	10.12
50	9.569442418479	0.106949822311	2.88e-06	11.27
55	9.569443486061	0.106950889894	1.81e-06	12.74
60	9.569444311557	0.106951715390	9.87e-07	13.95
65	9.569445012038	0.106952415871	2.87e-07	15.35
70	9.569445182508	0.106952586340	1.16e-07	16.67
75	9.569445225551	0.106952629384	7.34e-08	17.99
80	9.569445256661	0.106952660494	4.23e-08	19.49

Table 3.4: CDEO method compared with [ALO, Table 2]: Am. put, strike 100, spot $s = 100$, maturity $\vartheta = 1$. BS parameters: $r = q = 0.05, \sigma = 0.25$.

2.b: Next, we consider the Black-Scholes market with $r = q = 0.04$ and $\sigma = 0.2$. In [ALO, Subsection 6.1.2] different high precision algorithms from the literature are applied in order to calculate the value of an American put option with strike 100 and maturity $\vartheta = 3$ at the spot prices $s \in \{80, 100, 120\}$. We compute price approximations with the CDEO method and compare these to the most accurate approximation from [ALO, Table 3], i.e. the values contained in the column *True price*. To this end, we initialize Algorithm 3.18 with $n := 1200$ randomly generated basis functions of the type $f_k(x) := (a_k - e^x)_+$. The set Γ_1 is generated by independently drawing 400 samples from the distribution $\mathcal{L}(\mathcal{U}_{[0,3]}) \otimes \mathcal{L}(\log \mathcal{U}_{[50,100]})$. For the calculation of the CDEO we use the optimization point $(T, x_0) = (3, \log(100))$. The approximations to the American option value at the other two spot prices are obtained by simply evaluating the European value function associated to the CDEO. The columns $s = 80$, $s = 100$ and $s = 120$ contain the approximate option values generated by the CDEO method at the respective spot prices. The columns diff_{80}, diff_{100} and diff_{120} display the absolute difference between the CDEO approximation and the corresponding value from the column *True price* of [ALO, Table 3] for the different spot prices. Due to the fact that the values in [ALO, Table 3] are indicated with only five digits after the decimal point, we cannot compare the quality of the competing approximation methods beyond this precision. As above, the column *total time* indicates the total amount of seconds that has passed since the algorithm was started.

iteration	$s = 80$	diff_{80}	$s = 100$	diff_{100}	$s = 120$	diff_{120}	total time (s)
1	23.22730	0.00104	12.59989	0.00532	6.47783	0.00459	0.94
2	23.22644	0.00190	12.59994	0.00527	6.47778	0.00465	1.04
3	23.22809	0.00025	12.60152	0.00369	6.47981	0.00262	1.20
4	23.22883	0.00049	12.60348	0.00173	6.48139	0.00103	1.34
5	23.25013	0.02179	12.60425	0.00096	6.48195	0.00047	1.47
6	23.22657	0.00177	12.60468	0.00053	6.48210	0.00033	1.58
7	23.22777	0.00057	12.60492	0.00029	6.48217	0.00025	1.71
8	23.22795	0.00039	12.60492	0.00029	6.48217	0.00026	1.88
9	23.22792	0.00042	12.60493	0.00028	6.48218	0.00025	2.03
10	23.22822	0.00012	12.60493	0.00028	6.48217	0.00025	2.13
15	23.22830	0.00004	12.60514	0.00007	6.48236	0.00006	2.94
20	23.22827	0.00007	12.60516	0.00005	6.48239	0.00004	3.71
25	23.22833	0.00001	12.60518	0.00003	6.48240	0.00003	4.57
30	23.22834	0.00000	12.60521	0.00000	6.48242	0.00001	5.41
35	23.22834	0.00000	12.60521	0.00000	6.48242	0.00000	6.25

Table 3.5: CDEO method compared with [ALO, Table 3]: Am. put, strike 100, spot $s \in \{80, 100, 120\}$, maturity $\vartheta = 3$. BS parameters: $r = q = 0.04, \sigma = 0.2$.

In light of the numerical results from above, it is fair to say that the CDEO approach can compete with binomial tree and finite difference methods in view of accuracy and computational effort. Moreover, within the working accuracy of the optimization subroutines inherent in Algorithm 3.18, our price approximations are comparable to the high precision values from [ALO]. Similar to the binomial tree procedure, the Matlab implementation of Algorithm 3.18 is straightforward and requires only a few lines of source code. Conceptually and also from an implementation point of view, the algorithm put forward by [ALO] is more complicated than the CDEO approach. The same holds true for finite difference methods. Implementing an efficient and robust FDI-based pricing algorithm requires more prudence and ingenuity. For example, the choice of the discretization scheme, the construction of the grid and the boundary conditions at its borders can strongly influence the quality of the approximation - in particular if the derivative of the American payoff exhibits discontinuities. Moreover, in the case of the American put we were able to solve the finite dimensional linear complementary problems using the extremely fast Brennan Schwartz-Algorithm. The applicability of this direct method heavily depends on the structure of the American payoff, cf. [JLL]. In general, more complex and resource consuming algorithms are necessary in order to solve the LCPs arising as we iterate through the time layers, cf. [SD, Subsection 4.6.2].

As discussed above, the CDEO method generates a European payoff $\tilde{f}$ which allows us to easily calculate upper bounds for the American option value at any spot price and any maturity within the time horizon. Algorithm 3.18 compactly stores the CDEO payoff in a vector of some prespecified size, e.g. 1500 double values (= 12 kB) in *Test 1*. Our findings from Section 3.2 may be considered as indication that the American put is represented by its CDEO. Moreover, we observe that the European value function associated to $\tilde{f}$ indeed provides good approximations to the value of the American put within the associated continuation set. In addition, an estimate of the early exercise curve can be conveniently obtained from the mapping $v_{\mathrm{eu},\tilde{f}}$, cf. Theorem 2.52 and Section 3.4 below. In this sense Algorithm 3.18 can be considered as a global approximation method, whereas the classical binomial tree method and the approach from [ALO] generate single option prices. Clearly, finite difference schemes can be used to compute American option prices at any point of some grid in the price-time plane. Other option values may then be obtained by interpolation. Nevertheless, storing large grids is memory consuming. For example, a medium sized grid containing $10000 \times 10000 = 10^8$ double precision gridpoints already requires 800 MB of memory. In order to enhance the accuracy of the price approximations, it is necessary to increase the resolution of the grid. In the setting of *Test 1* we required to use a grid of the size 5.29×10^{10} in order that the finite difference method produces an option value that coincides to 7 significant digits with the CDEO approximation, see Table 3.3. Storing the complete grid containing 5.29×10^{10} double precision values requires 423.2 GB of memory.
In the setting of *Test 2.a* the finite difference method does not perform better. Table 1 from [ALO] indicates that a grid of the size $250000 \times 250000 = 6.25 \times 10^{10}$ is necessary such that the finite difference approximation to the American premium coincides up to 7 significant digits with the CDEO ($m = 80$) value from Table 3.4. Storing a complete double precision grid of this size requires about 500 GB, whereas storing the corresponding approximate CDEO requires 14.4 kB of memory. The authors of [ALO] state that the numerical results from [ALO, Table 1] were obtained using “a production-quality

Crank-Nicolson finite difference method". Consequently, it might be worthwhile to consider the CDEO approach as a useful supplement to the algorithm toolbox for American option pricing.

3.4 Lower bounds and estimation of the early exercise curve

Consider the univariate Black-Scholes market (1.3) and assume until the end of this section that the Brownian motion W is supported on a right-continuous, complete filtration. Furthermore, suppose that $g : \mathbb{R} \to \mathbb{R}_+$ is a continuous American payoff satisfying the growth condition (3.10). We reuse the notation (1.4, 1.6) from Section 1.2 for European and American value functions. In Section 3.1 we discussed how approximations to a cheapest dominating European payoff can be obtained by solving a semi-infinite linear program of the type $\mathrm{P}_n^{\mathrm{c}}$ numerically. Proposition 2.3 warrants that the European value function associated to any solution of $\mathrm{P}_n^{\mathrm{c}}$ always constitutes a global upper bound of the mapping $v_{\mathrm{am},g}$.

In case that g is representable on $C_{(T,x_0)}$ by some measure μ^*, the latter coincides with the generalized CDEO of g with respect to T, x_0, cf. Lemma 3.2. Moreover, Lemma 3.7 indicates that solving the optimization task $\mathrm{P}_n^{\mathrm{c}}$ may indeed yield a reasonable approximation to μ^* if the discretization was chosen adequately and the number of basis elements n is sufficiently large. In particular, the corresponding European value function should provide tight upper bounds for the American value function $v_{\mathrm{am},g}$ on the set $C_{(T,x_0)}$. Above, we applied the CDEO procedure to the American put in the Black-Scholes market for illustrative and benchmarking purposes. We observed that the price surface associated to approximate CDEO exhibits all the qualitative features which are required by Theorem 2.52 in order to warrant representability, cf. Section 3.2. Affirmatively, within the working accuracy of the algorithm, the option prices stemming from the CDEO method complied with high-precision approximations from the literature, cf. Table 3.4.

This subsection aims at presenting a duality method which provides us with lower bounds for the American value function - no matter whether or not the American payoff is representable. The central idea of the approach is to construct approximations to the relevant sections of the early exercise boundary which take the position information encoded within the D_{Γ_m}-optimal elements from Algorithm 3.18 into account. Irrespective of representability issues, the practitioner may combine the CDEO procedure with the algorithm from this section in order to compute upper and lower bounds for American option prices in a handy manner. Clearly, the discussion below can be generalized in many ways. From our point of view, this would require a more cumbersome notation which might obscure the basic mathematical principles. We invite the interested reader to refine the discussion below according to her requirements.

The method is based on the following simple observation: Suppose that f is a European payoff that superreplicates g up to some terminal time $T \in \mathbb{R}_{++}$. For any Borel set $M \in \mathcal{B}([0,T] \times \mathbb{R})$ and any $\vartheta \in \mathbb{R}_+$ we define the first entry time

$$\tau_\vartheta^M := \inf\{t \in \mathbb{R}_+ \mid (\vartheta - t, X_t) \in M\} \wedge \vartheta.$$

Owing to the fact that the filtration of the Brownian motion is assumed to be right-continuous and complete, the Début theorem warrants that τ_ϑ^M is a stopping time, cf. [KB, Theorem 7.7]. In line with (2.17), we denote by

$$H_M(\vartheta, x) := \mathbb{P}_x\left((\vartheta - \tau_\vartheta^M, X_{\tau_\vartheta^M}) \in M\right) \tag{3.21}$$

the probability that the space-time process started at time $T - \vartheta$ and log-price x hits the set M. In case that $H_M(\vartheta, x) = 1$ holds true, we obtain

$$\begin{aligned} v_{\mathrm{am},g}(\vartheta, x) &\geq \mathbb{E}_x\left[e^{-r\tau_\vartheta^M}\, g(X_{\tau_\vartheta^M})\right] \\ &\geq \mathbb{E}_x\left[e^{-r\tau_\vartheta^M}\, v_{\mathrm{eu},f}(\vartheta - \tau_\vartheta^M, X_{\tau_\vartheta^M})\right] - \sup_{(\vartheta', x') \in M}\left\{v_{\mathrm{eu},f}(\vartheta', x') - g(x')\right\} \\ &= v_{\mathrm{eu},f}(\vartheta, x) - \sup_{(\vartheta', x') \in M}\left\{v_{\mathrm{eu},f}(\vartheta', x') - g(x')\right\}. \end{aligned} \tag{3.22}$$

The last equality follows by applying the optional sampling theorem to the discounted European value process. Let us remark that the conditions imposed on the payoff g can be relaxed without affecting the validity of the latter arguments. More specifically, it suffices to assume that g is an upper semi-continuous function satisfying the integrability condition (1.5). Maximizing the right-hand side of inequality (3.22) over all Borel sets $M \in \mathcal{B}([0,T] \times \mathbb{R})$ satisfying $H_M(\vartheta, x) = 1$, directly yields the following proposition.

3.23 Proposition: *Let $g : \mathbb{R} \to \mathbb{R}_+$ denote an upper semi-continuous American payoff meeting the integrability condition* (1.5)*. Furthermore, suppose that $f : \mathbb{R} \to \mathbb{R}_+$ is a European claim which superreplicates g up to time $T \in \mathbb{R}_{++}$. For any $(\vartheta, x) \in [0,T] \times \mathbb{R}$ we have*

$$0 \leq v_{\mathrm{eu},f}(\vartheta, x) - v_{\mathrm{am},g}(\vartheta, x) \leq \inf_{\substack{M \in \mathcal{B}([0,T] \times \mathbb{R}) \\ H_M(\vartheta, x) = 1}} \sup_{(\vartheta', x') \in M}\left\{v_{\mathrm{eu},f}(\vartheta', x') - g(x')\right\}. \tag{3.24}$$

The reader easily verifies that in case of representability the latter result blends in with our findings from Subsection 2.1.1. In particular, suppose that the set M is defined as in Equation (2.13). For any $(\vartheta, x) \in [0,T] \times \mathbb{R}$ with $H_M(\vartheta, x) = 1$, we obtain from (3.24) that

$$0 \leq v_{\mathrm{eu},f}(\vartheta, x) - v_{\mathrm{am},g}(\vartheta, x) \leq \sup_{(\vartheta', x') \in M}\left\{v_{\mathrm{eu},f}(\vartheta', x') - g(x')\right\} = 0$$

holds true. This statement corresponds to the first assertion of Proposition 2.16. Moreover, the reader may gain more insight on the interplay between (3.24) and the notion of local representability by applying Proposition 3.23 within the setting of Example 2.25.

More importantly, no matter whether or not the American payoff g is representable, Equation (3.24) provides the practitioner with a tool to obtain upper and lower bounds for the associated value function $v_{\mathrm{am},g}$. The concrete implementation of (3.24) involves three central subtasks:

1. Computing a superreplicating European option f.
2. Generating some set M which satisfies $H_M(\vartheta, x) = 1$ for any maturity/log-price pair (ϑ, x) of interest.
3. Determine $\sup_{(\vartheta',x')\in M}\{v_{\mathrm{eu},f}(\vartheta', x') - g(x')\}$ numerically.

As we desire our procedure to be compatible with the concept of representability, the canonic choice for the superreplicating payoff from *Subtask 1* is the cheapest dominating European option with respect to some suitably chosen optimization point. Under certain regularity conditions, see Proposition 3.4 and Lemma 3.15, the latter can be approximated by numerically solving a semi-infinite linear program as outlined above.
Let us elaborate on some details of a possible approach: Suppose that we desire to approximate the American option price at some point $(T, x_0) \in \mathbb{R}_{++} \times \mathbb{R}$ which is located within in the continuation region associated to g. Choose a set $f_1, f_2, ...$ of European payoff functions satisfying the requirements from Lemma 3.15. The latter warrants that the optimal value in Program $\mathrm{P}_n^{\mathrm{c}}$ is attained by some vector $a^{*,n} \in \mathbb{R}_+^n$ and that there is no duality gap. Moreover, the dual optimal value is attained by some discrete measure λ_n^* concentrated on $(\vartheta_i^n, x_i^n)_{i=1,\dots,n+1} \subset [0, T] \times \mathbb{R}$. In order to shorten the notation, we write

$$f^{*,n} := \textstyle\sum_{k=1}^n a_k^{*,n} f_k$$

and

$$\Psi_n := v_{\mathrm{eu},f^{*,n}} - g.$$

Clearly, the European payoff $f^{*,n}$ superreplicates g up to time T which yields that $v_{\mathrm{eu},f^{*,n}} \geq v_{\mathrm{am},g} \geq g$ and therefore $\Psi_n \geq 0$ holds true on $[0, T] \times \mathbb{R}$, cf. Proposition 2.3. In light of Equation (3.24), we desire to construct a Borel set $M \in \mathcal{B}([0, T] \times \mathbb{R})$ such that the discrepancy

$$\Delta_n(M) := \sup_{(\vartheta',x')\in M} \Psi_n(\vartheta', x')$$

is small and such that $H_M(T, x_0) = 1$. Given any $M \in \mathcal{B}([0, T] \times \mathbb{R})$, we denote by

$$d_n(M) := \sup_{(\vartheta',x')\in M} \inf_{\substack{(\vartheta,x)\in[0,T]\times\mathbb{R} \\ \Psi_n(\vartheta,x)=0}} \|(\vartheta, x) - (\vartheta', x')\|_2 \tag{3.25}$$

the asymmetric Hausdorff distance from M to the zero set of Ψ_n. The mass of the dual optimizer provides us with some information concerning the zeros of the mapping Ψ_n. Indeed, owing to the first complementary slackness condition from (3.11), we have $\Psi_n(\vartheta_i^n, x_i^n) = 0$ for any $i \in \{1, .., n + 1\}$ such that $\lambda_n^*(\vartheta_i^n, x_i^n) > 0$. In particular, we can conclude that the set $\{\Psi_n = 0\}$ is not empty and therefore $d_n(M)$ is finite. Now, assume that Ψ_n admits a modulus of continuity ω_n on $[0, T] \times \mathbb{R}$ or at least on some suitably chosen subset S containing M. In other words, there exists a continuous, non-decreasing function $\omega_n : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\omega_n(0) = 0$ and

$$|\Psi_n(\vartheta', x') - \Psi_n(\vartheta, x)| \leq \omega_n(\|(\vartheta', x') - (\vartheta, x)\|_2)$$

holds true for all $(\vartheta', x'), (\vartheta, x) \in S$. From (3.25) it is apparent that for any $\varepsilon > 0$ and any $(\vartheta', x') \in M$ there exists some $(\vartheta_{\vartheta'}, x_{x'}) \in [0, T] \times \mathbb{R}$ satisfying $\Psi_n(\vartheta_{\vartheta'}, x_{x'}) = 0$ such that $\|(\vartheta', x') - (\vartheta_{\vartheta'}, x_{x'})\|_2 \leq d_n(M) + \varepsilon$. Due to the fact that ω_n is non-decreasing, we obtain

$$\Delta_n(M) = \sup_{(\vartheta', x') \in M} |\Psi_n(\vartheta', x') - \Psi_n(\vartheta_{\vartheta'}, x_{x'})| \leq \omega_n(d_n(M) + \varepsilon).$$

Taking the limit $\varepsilon \searrow 0$ yields $\Delta_n(M) \leq \omega_n(d_n(M))$. In light of (3.24), we can finally conclude that

$$\sup_{\substack{(\vartheta, x) \in [0,T] \times \mathbb{R} \\ H_M(\vartheta, x) = 1}} |v_{\mathrm{eu}, f^{*,n}}(\vartheta, x) - v_{\mathrm{am}, g}(\vartheta, x)| \leq \Delta_n(M) \leq \omega_n(d_n(M)). \tag{3.26}$$

The latter equation reflects the idea that the quality of the European approximation to the American option price at $(\vartheta, x) \in [0, T] \times \mathbb{R}$ depends in some sense continuously on the approximability of the null set of the mapping Ψ_n by Borel sets M which are almost surely hit by the space-time process started at $(T - \vartheta, x)$, i.e. $H_M(\vartheta, x) = 1$. From Proposition 2.16 we know that $\{\Psi_n = 0\}$ is a subset of the stopping region associated to g. In applications we might encounter the following situation: On the grounds of some preceding considerations, we suspect that the American payoff g is represented with respect to T, x_0 by some measure μ^*. We choose a discretization which we deem reasonable to approximate the representing measure in the sense of Lemma 3.7. If true, the latter warrants that a subsequence of optimizers associated to the finite dimensional subproblems converges weakly to μ^*. The finite dimensional surrogates can be numerically solved by applying Algorithm 3.18. By further thinning out the subsequence, we can achieve that the associated dual optimizers converge vaguely to some measure λ^*_∞ satisfying the slackness condition (3.16), cf. Lemma 3.15. Equation (3.16) warrants that the support of λ^*_∞ is located within the stopping set associated to g. Besides, if for any $x \in \mathbb{R}$ the mapping $(0, T) \ni \vartheta \mapsto v_{\mathrm{eu}, \mu^*}(\vartheta, x)$ assumes its minimal value at some uniquely determined maturity, we can conclude that the support of the measure λ^*_∞ is located on the early exercise curve associated to g. Our numerical findings from Section 3.2 indicated that this situation indeed seems to occur for the American put.

Consequently, we propose the following heuristic: Use the position information of the dual mass in order to construct an approximation to the early exercise boundary associated to g. In other words, use the dual optimizer obtained by Algorithm 3.18 in order to find a candidate set M which yields a valuable error bound in Proposition 3.23.

For illustrative purposes, let us discuss a possible implementation of this heuristic for the American put in the Black-Scholes market (1.3). Please, note that we will use non-logarithmic prices until the end of this section. First, consider the setting of *Test 1* from Section 3.3, i.e. the model parameters are given by $r = 0.06, \sigma = 0.4, T = 0.5$, there are no dividend payments and we have a put option with strike $K = 100$. The basis functions, the set of initial constraints and also the other parameters for Algorithm 3.18 are chosen as described in Test 1, cf. page 87. Let $\tilde{f}$ denote a solution of $\mathrm{P}_n^{\mathrm{c}}$ generated by Algorithm 3.18. We fit a smoothing spline γ to the support of the D_{Γ_m}-optimal measure λ^m from the last ($m_{\max} = 100$) iteration of Algorithm 3.18. For an introduction to smoothing spline techniques we refer the reader to [WB]. The support of λ^m, the spline γ and an approximation to the early exercise boundary obtained by a finite difference method on a fine grid are displayed in Figure 3.6. In line with (3.21), we denote by $H_{b(\gamma)}(\vartheta, s)$ the probability that the space-time process started at time $T - \vartheta$ and price s hits the set

$$b(\gamma) := \{(\vartheta', s') \in [0, T] \times \mathbb{R}_{++} \mid s' = \gamma(\vartheta') \quad \text{or} \quad \vartheta' = 0, s' \geq \gamma(0)\}.$$

The latter serves as approximation to the early exercise boundary of the put.

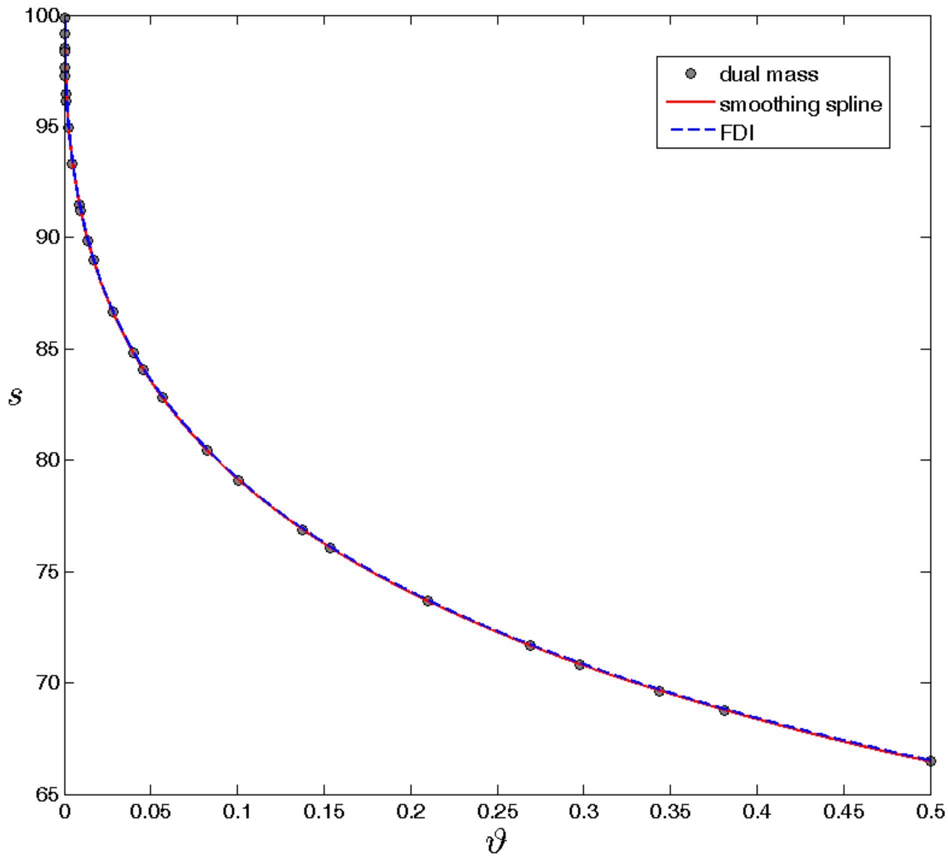

Figure 3.6: Smoothing spline and FDI approximation to the early exercise boundary of the American put, setting of Test 1 from Section 3.3.

Clearly, for any $(\vartheta, s) \in (0, T] \times \mathbb{R}_{++}$ such that $s > \gamma(\vartheta)$ we have $H_{b(\gamma)}(\vartheta, s) = 1$ and therefore Proposition 3.23 yields

$$0 \leq v_{\mathrm{eu},\tilde{f}}(\vartheta, s) - v_{\mathrm{am},g}(\vartheta, s) \leq \sup_{\vartheta' \in (0,T]} \left\{ v_{\mathrm{eu},\tilde{f}}(\vartheta', \gamma(\vartheta')) - g(\gamma(\vartheta')) \right\} =: \Delta(\gamma). \tag{3.27}$$

We compare the upper and lower bounds for $v_{\mathrm{am},g}(0.5, 100)$ from (3.27) with a price approximation obtained by a Crank-Nicolson finite difference scheme which required about the same amount of computation time, cf. Table 3.6.

$$\Delta(\gamma) = 0.000029694563$$

	$v_{\mathrm{am},g}(0.5, 100)$
CDEO upper bound	9.945136132710
CDEO lower bound	9.945106438147
FDI approximation	9.945135392820

Table 3.6: CDEO upper and lower bounds for $v_{\mathrm{am},g}(0.5, 100)$, setting of Test 1 from Section 3.3.

Analogously, we apply the procedure from above within the setting of *Test 2.a* from Section 3.3, i.e. we consider an American put with strike 100 maturing at $T = 1$ in the Black-Scholes market with $r = 0.05, \sigma = 0.25$ and a continuous dividend rate of $q = 0.05$. Here, we choose to terminate Algorithm 3.18 after $m_{\max} = 100$ iterations. The other input parameters, in particular the basis functions and the initial set of constraints, are chosen as described in Test 2.a, cf. page 90. The smoothing spline approximation to the early exercise boundary is based on the D_{Γ_m}-optimal measure from the last iteration of the algorithm, cf. Figure 3.7. The upper and lower bounds obtained by the CDEO method are compared against the high-precision value $(l, m, n) = (65, 8, 32)$ from [ALO, Table 2]. In addition, a finite difference price approximation which required about the same amount of computation time is listed in Table 3.7. The column *diff* indicates the deviation from the [ALO] reference premium.

$$\Delta(\gamma) = 0.000055146019$$

	Am. premium	diff
[ALO, Table 2]	0.106952702747	0
CDEO upper bound	0.106952689003	1.37e-08
CDEO lower bound	0.106897542984	5.52e-05
FDI approximation	0.106945552058	7.15e-06

Table 3.7: CDEO upper and lower bounds for the American premium, setting of Test 2.a from Section 3.3.

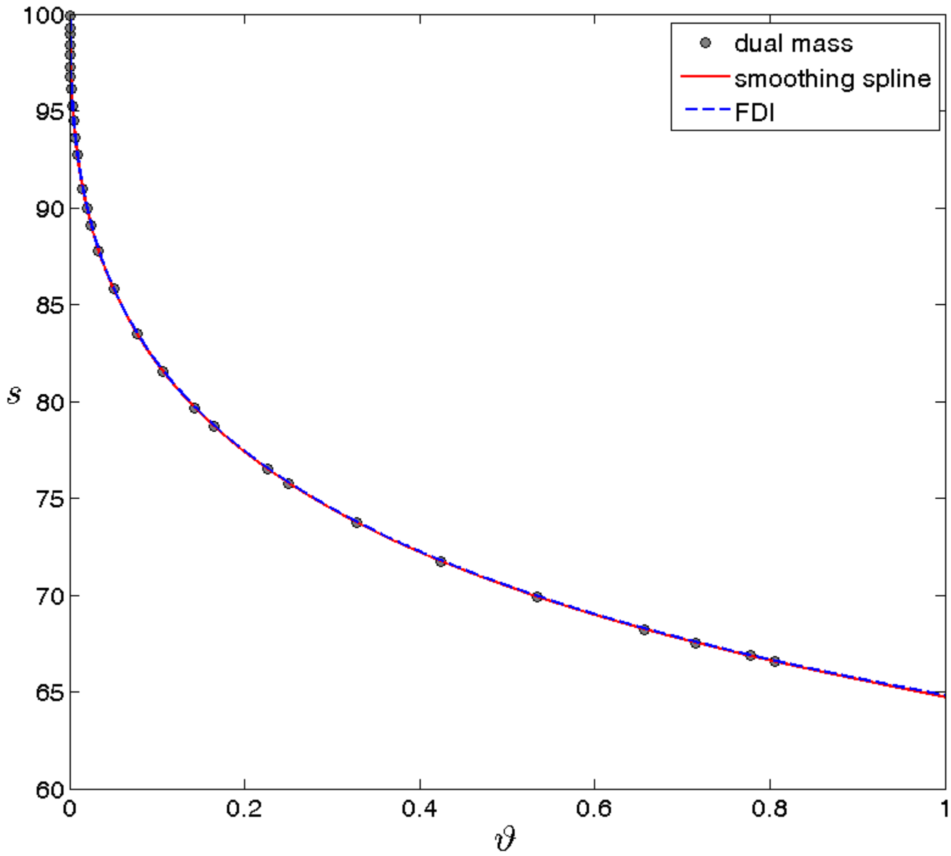

Figure 3.7: Smoothing spline and FDI approximation to the early exercise boundary of the American put, setting of Test 2.a from Section 3.3.

Finally, we assess the quality of the lower bounds obtained by the CDEO method within the setting of *Test 2.b* from Section 3.3, i.e. we consider a Black-Scholes market with $r = q = 0.04, \sigma = 0.2$ and an American put option with strike $K = 100$ and maturity $T = 3$. Algorithm 3.18 is terminated after $m_{\max} = 100$ iterations. Any other input parameter is chosen according to the description from Test 2.b, cf. page 91. For the computation of the CDEO we use the optimization point $(T, x_0) = (3, \log(100))$. Figure 3.8 compares a finite difference approximation to the exercise curve with a smoothing spline approximation which is based on the dual optimizer from the last iteration. Table 3.8 contains the CDEO upper and lower bounds at the different spot prices as well as the corresponding values from the column *True price* of [ALO, Table 3]. Due to the fact that the latter are indicated with only five digits after the decimal point, we cannot compare the quality of the competing approximations beyond this precision. Within this numerical accuracy the CDEO upper and lower bounds coincide with the reference values from [ALO, Table 3].

In conclusion, we find that in all three test cases the CDEO method provides us with tight bounds for the American option value and a numerically convenient spline approximation to the early exercise boundary. Moreover, the basis vector associated to the approximate CDEO and the spline coefficients can be compactly stored using only a few kilobytes of memory. As argued in Section 3.3, the parsimonious nature of the CDEO method is in particular useful when global price approximations are desired. Let us remark that the European value function associated to the CDEO as well as the smoothing spline γ can be evaluated with very little computational effort.

$$\Delta(\gamma) = 0.000002089720$$

	$v_{\mathrm{am},g}(3, 80)$	$v_{\mathrm{am},g}(3, 100)$	$v_{\mathrm{am},g}(3, 120)$
[ALO, Table 3]	23.22834	12.60521	6.48242
CDEO upper bound	23.22834	12.60521	6.48242
CDEO lower bound	23.22834	12.60521	6.48242

Table 3.8: CDEO upper and lower bounds for $v_{\mathrm{am},g}(3, s), s \in \{80, 100, 120\}$, setting of Test 2.b from Section 3.3.

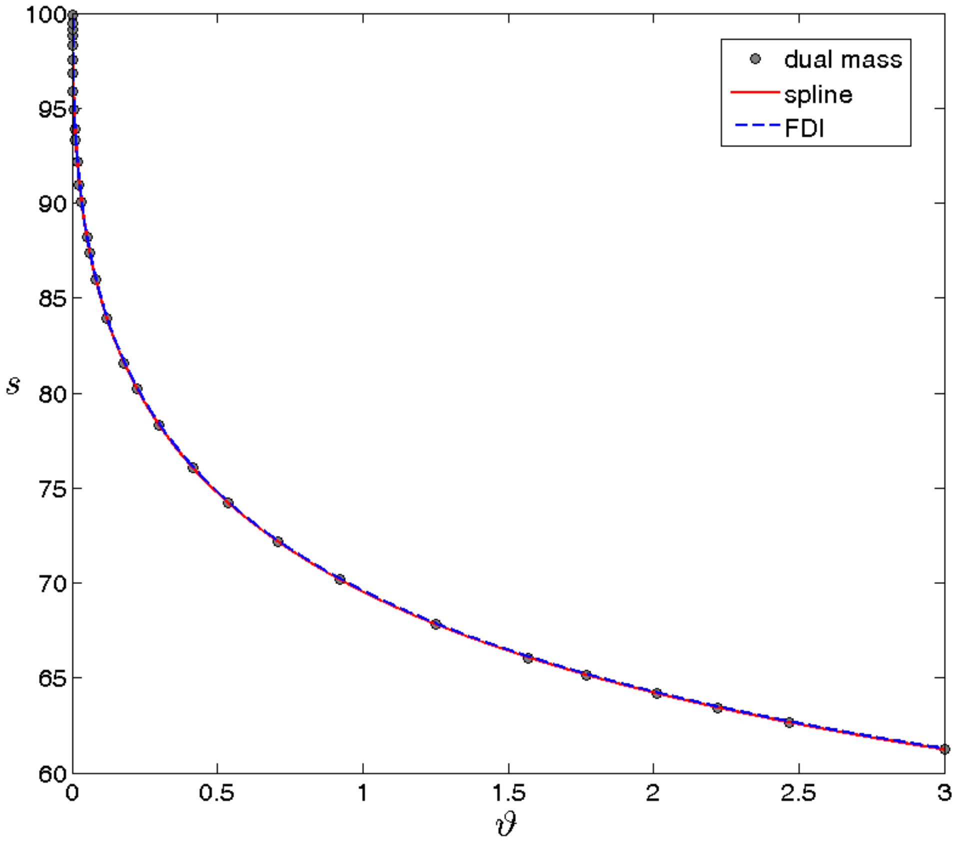

Figure 3.8: Smoothing spline and FDI approximation to the early exercise boundary of the American put, setting of Test 2.b from Section 3.3.

4 Conclusion

The thesis at hand aims at analyzing and linking together the mathematical notion of representability, embedded American payoffs and cheapest dominating European options. For the ease of exposition, we choose the univariate Black-Scholes market. It is argued that CDEOs are in some sense inversely related to the embedding operation. This process reveals a new duality structure between European and American valuation problems which we deem as very fruitful for future research. As a by-product, we derive some non-trivial statements concerning free boundary problems and a Markovian-style super martingale decomposition. Furthermore, we demonstrate that it is reasonable to understand representability as a local property of the connected components associated to the continuation region of the American claim at hand. By studying the American option embedded into the European put within the Black-Scholes market, we conclude that American options may be representable up to some maximal time horizon. In addition, we provide several other explicit examples of representable American claims and study their analyticity and smoothness properties. Relying on methods from infinite-dimensional optimization, we make a first step towards verifying representability of certain American claims. The results of Section 3.2 suggest in particular that representability holds for the prime example of an American put in the Black-Scholes model, contrary to the indications following from the analysis in [JM2]. This gives new hope that the original endeavor of Jourdain and Martini may ultimately lead to a positive answer and that their concept of embedded American options has a broader scope than expected. Moreover, we discuss some computational aspects related to the CDEO algorithm which generates upper bounds for American option prices, regardless of any representability issues. Based on the Lagrange dual associated to the CDEO optimization task, we propose a new method which allows us to construct spline approximations to the early exercise curve and generate lower bounds for American option prices. For the American put in the Black-Scholes market, these upper and lower bounds are benchmarked against high-precision methods from the literature.

As an ambitious goal for future research it remains to fully characterize representability of American options in the Black-Scholes model and more general markets driven by univariate or multivariate continuous diffusion processes. In particular, a rigorous proof for the American put is still wanting. Let us remark that this thesis is formulated in classical terms from mathematical finance. Clearly, the notions from above are easily translated into the language of Markov process theory. Besides, throughout the thesis the interested reader can find several remarks concerning possible extensions.

5 Supplementary material

This chapter compiles a small collection of mathematical tools. In particular, we give a very brief summary of some basic notions from functional analysis and convex optimization.

5.1 Lemmata

The following lemmata are rather elementary. Short proofs are provided in order to make the thesis at hand more self-contained.

5.1 Lemma: *Let $I \subset \mathbb{R}$ denote an open interval containing the origin and suppose that $f : I \to [0, +\infty)$ is a non-negative function with $f(0) = 0$. If f can be continued analytically to some open set containing the origin, then there exists an $\varepsilon > 0$ such that $f''(x) \geq 0$ for any $x \in (-\varepsilon, \varepsilon)$.*

Proof. Assume that the assertion of the lemma is false, i.e. for any $\varepsilon > 0$ exists some point $x \in (-\varepsilon, \varepsilon)$ such that $f''(x) < 0$. Due to the fact that f assumes its minimal value at 0, we have $f'(0) = 0$ and $f''(0) \geq 0$. Consequently, the following three cases can be directly excluded:

1. $\exists \varepsilon > 0$ such that $f''(x) < 0$ for all $x \in (-\varepsilon, \varepsilon) \setminus \{0\}$.
2. $\exists \varepsilon > 0$ such that $f''(x) < 0$ for all $x \in (-\varepsilon, 0)$ and $f''(x) \geq 0$ for all $x \in (0, \varepsilon)$.
3. $\exists \varepsilon > 0$ such that $f''(x) \geq 0$ for all $x \in (-\varepsilon, 0)$ and $f''(x) < 0$ for all $x \in (0, \varepsilon)$.

Hence, there remain only two possibilities:

4. $\forall \varepsilon > 0$ exist $y^+, y^- \in (-\varepsilon, 0)$ such that $f''(y^+) \geq 0$ and $f''(y^-) < 0$.
5. $\forall \varepsilon > 0$ exist $y^+, y^- \in (0, \varepsilon)$ such that $f''(y^+) \geq 0$ and $f''(y^-) < 0$.

In case that the fourth statement holds true, we can find sequences $y_n^+, y_n^- \nearrow 0$ as $n \to \infty$ with $\max\{y_{n-1}^+, y_{n-1}^-\} < \min\{y_n^+, y_n^-\}$ and $f''(y_n^-) < 0 \leq f''(y_n^+)$. The mean value theorem implies that there exists a sequence $z_n \nearrow 0$ such that $f'''(z_n) = 0$. Due to the fact that the function f and its derivatives are analytic on some open ball B containing the origin we, find that $f'''(z) = 0$ for all $z \in B$. Hence f'' is constant on some interval containing the origin which clearly contradicts your assumption. Along the same lines the reader can verify that the fifth statement yields a contradiction as well. This finally proves the lemma at hand. $\square$

5.2 Lemma: *Let* $y \mapsto \mathrm{N}(\mu, \sigma^2, y)$ *denote the probability density function of a normal distribution with mean* μ *and variance* σ^2*. Then:*

1. *For any* $y \in \mathbb{R}, \mu, \sigma^2 \in \mathbb{C}$ *with* $\operatorname{Re}\sigma^2 > 0$ *have*
$$|\mathrm{N}(\mu, \sigma^2, y)| = \frac{\exp\left(-\frac{\operatorname{Re}\sigma^2}{2|\sigma^2|^2}\left(y - \operatorname{Re}\mu - \frac{\operatorname{Im}\mu \operatorname{Im}\sigma^2}{\operatorname{Re}\sigma^2}\right)^2 + \frac{(\operatorname{Im}\mu)^2}{2\operatorname{Re}\sigma^2}\right)}{\sqrt{2\pi|\sigma^2|}}.$$

2. *For any* $\mu, \tilde{\mu} \in \mathbb{C}$ *and* $\sigma, \tilde{\sigma} \in \mathbb{C} \setminus \{0\}$ *with* $\sigma \neq \tilde{\sigma}$ *we have*
$$\frac{\mathrm{N}(\mu, \sigma^2, y)}{\mathrm{N}(\tilde{\mu}, \tilde{\sigma}^2, y)} = \frac{\tilde{\sigma}}{\sigma} \exp\left(-\frac{(y-A)^2}{2B}\right) \exp\left(-\frac{(\mu - \tilde{\mu})^2}{2(\sigma^2 - \tilde{\sigma}^2)}\right)$$
where the quantities A *and* B *are defined as follows:*
$$A := \frac{\tilde{\mu}\sigma^2 - \mu\tilde{\sigma}^2}{\sigma^2 - \tilde{\sigma}^2} \qquad B := \frac{\tilde{\sigma}^2\sigma^2}{\tilde{\sigma}^2 - \sigma^2}$$

3. *For any* $\mu, \tilde{\mu} \in \mathbb{C}$ *and* $\sigma \in \mathbb{C} \setminus \{0\}$ *we have*
$$\frac{\mathrm{N}(\mu, \sigma^2, y)}{\mathrm{N}(\tilde{\mu}, \sigma^2, y)} = \exp\left(y\,\frac{\mu - \tilde{\mu}}{\sigma^2}\right) \exp\left(\frac{\tilde{\mu}^2 - \mu^2}{2\sigma^2}\right).$$

Proof.

1. For any $y \in \mathbb{R}$ and $\mu, \sigma^2 \in \mathbb{C}$ such that $\operatorname{Re}\sigma^2 > 0$ we obtain
$$\begin{aligned}
&\left|\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)\right| \\
&= \exp\left(-\operatorname{Re}\frac{(y-\mu)^2}{2\sigma^2}\right) \\
&= \exp\left(-\frac{\operatorname{Re}\sigma^2 \operatorname{Re}((y-\mu)^2) + \operatorname{Im}\sigma^2 \operatorname{Im}((y-\mu)^2)}{2|\sigma^2|^2}\right) \\
&= \exp\left(-\frac{(y - \operatorname{Re}\mu)^2 \operatorname{Re}\sigma^2 - 2(y - \operatorname{Re}\mu)\operatorname{Im}\sigma^2 \operatorname{Im}\mu - \operatorname{Re}\sigma^2(\operatorname{Im}\mu)^2}{2|\sigma^2|^2}\right).
\end{aligned}$$
Completing the square in the variable y yields
$$\begin{aligned}
&\left|\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)\right| \\
&= \exp\left(-\frac{\operatorname{Re}\sigma^2}{2|\sigma^2|^2}\left(y - \operatorname{Re}\mu - \frac{\operatorname{Im}\mu \operatorname{Im}\sigma^2}{\operatorname{Re}\sigma^2}\right)^2 + \frac{(\operatorname{Im}\mu)^2}{2|\sigma^2|^2}\left(\operatorname{Re}\sigma^2 + \frac{(\operatorname{Im}\sigma^2)^2}{\operatorname{Re}\sigma^2}\right)\right) \\
&= \exp\left(-\frac{\operatorname{Re}\sigma^2}{2|\sigma^2|^2}\left(y - \operatorname{Re}\mu - \frac{\operatorname{Im}\mu \operatorname{Im}\sigma^2}{\operatorname{Re}\sigma^2}\right)^2 + \frac{(\operatorname{Im}\mu)^2}{2\operatorname{Re}\sigma^2}\right)
\end{aligned}$$
and this proves the assertion.

2. Simple algebraic manipulations show that

$$\begin{aligned}\exp\left(\frac{1}{2}\left(\frac{y-\tilde{\mu}}{\tilde{\sigma}}\right)^2-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2\right)&=\exp\left(y^2\frac{\sigma^2-\tilde{\sigma}^2}{2\tilde{\sigma}^2\sigma^2}-y\frac{\tilde{\mu}\sigma^2-\mu\tilde{\sigma}^2}{\tilde{\sigma}^2\sigma^2}+\frac{\tilde{\mu}^2\sigma^2-\mu^2\tilde{\sigma}^2}{2\tilde{\sigma}^2\sigma^2}\right)\\&=\exp\left(-\frac{(y-A)^2}{2B}+\frac{A^2}{2B}+\frac{\tilde{\mu}^2\sigma^2-\mu^2\tilde{\sigma}^2}{2\tilde{\sigma}^2\sigma^2}\right)\\&=\exp\left(-\frac{(y-A)^2}{2B}\right)\exp\left(-\frac{(\tilde{\mu}-\mu)^2}{2(\sigma^2-\tilde{\sigma}^2)}\right).\end{aligned}$$

3. This is obvious.

□

The following lemma concerns generalized European value functions within the scope of the Black-Scholes market (2.41).

5.3 Lemma: *For $\hat{r}\in\mathbb{R},\sigma>0$ and some measure $\mu\in\mathcal{M}^+(\mathbb{R})$, we define the generalized European value function*

$$V(\vartheta,x):=\int_{-\infty}^{\infty}\mathrm{N}\left(x+\hat{r}\vartheta,\sigma^2\vartheta,y\right)\mathrm{d}\mu(y).$$

Suppose there exists some $(T,x_0)\in\mathbb{R}_{++}\times\mathbb{R}$ such that $V(T,x_0)<\infty$. Then, the mapping V is analytic on the open $\mathbb{C}^2$-domain

$$E:=\left\{\vartheta\in\mathbb{C}\;\middle|\;\sqrt{(\operatorname{Re}\vartheta-T/2)^2+(\operatorname{Im}\vartheta)^2}<T/2\right\}\times\mathbb{C}.$$

Proof. Due to the assumption $V(T,x_0)<\infty$, we find that

$$\mathrm{d}\nu(y):=\mathrm{N}\left(x_0+\hat{r}T,\sigma^2T,y\right)\mathrm{d}\mu(y)$$

is a finite measure. Applying Lemma 5.2 yields

$$\begin{aligned}V(\vartheta,x)&=\int_{-\infty}^{\infty}\frac{\mathrm{N}\left(x+\hat{r}\vartheta,\sigma^2\vartheta,y\right)}{\mathrm{N}\left(x_0+\hat{r}T,\sigma^2T,y\right)}\mathrm{d}\nu(y)\\&=\sqrt{\frac{T}{\vartheta}}\exp\left(\frac{\left(x_0-x+\hat{r}(T-\vartheta)\right)^2}{2\sigma^2(T-\vartheta)}\right)\int_{-\infty}^{\infty}\exp\left(-\frac{(y-A(\vartheta,x))^2}{2B(\vartheta,x)}\right)\mathrm{d}\nu(y)\end{aligned}$$

where $A:=A(\vartheta,x):=\frac{xT-x_0\vartheta}{T-\vartheta}$ and $B:=B(\vartheta,x):=\sigma^2\frac{\vartheta T}{T-\vartheta}$. Consequently, we only need to show that the function

$$F(\vartheta,x):=\int_{-\infty}^{\infty}\exp\left(-\frac{(y-A(\vartheta,x))^2}{2B(\vartheta,x)}\right)\mathrm{d}\nu(y)$$

is analytic on E. Owing to Hartogs' theorem, see [KR, Paragraph 2.4], it is enough to verify that F is partially analytic on E. Lemma 5.2 yields

$$\left|\exp\left(-\frac{(y-A)^2}{2B}\right)\right|=|h(\vartheta,x)|\exp\left(-\frac{\operatorname{Re}B}{2|B|^2}\left(y-\operatorname{Re}A-\frac{\operatorname{Im}A\operatorname{Im}B}{\operatorname{Re}B}\right)^2\right)$$

for all $(\vartheta, x) \in E$ and $y \in \mathbb{R}$ where h denotes a certain mapping which is continuous on E. Moreover, for any $(\vartheta, x) \in E$ we have

$$\operatorname{Re} B = \sigma^2 T \frac{\operatorname{Re}\vartheta(T - \operatorname{Re}\vartheta) - (\operatorname{Im}\vartheta)^2}{|T-\vartheta|^2} > 0$$

and therefore

$$\sup_{y\in\mathbb{R}} \left| \exp\left(-\frac{(y-A)^2}{2B}\right)\right| \leq |h(\vartheta, x)| \tag{5.4}$$

is certainly bounded on every compact subset of E. For any closed contour γ contained in $\pi_1(E) := \{\vartheta \in \mathbb{C} \mid (\operatorname{Re}\vartheta - T/2)^2 + (\operatorname{Im}\vartheta)^2 < T^2/4\}$ and any $x \in \mathbb{C}$, we find that

$$\oint_\gamma F(\vartheta, x)\mathrm{d}\vartheta = \int_{-\infty}^{\infty} \oint_\gamma \exp\left(-\frac{(y - A(\vartheta, x))^2}{2B(\vartheta, x)}\right) \mathrm{d}\vartheta \mathrm{d}\nu(y) = 0.$$

In virtue of Morera's theorem, we conclude that the function $\vartheta \mapsto F(\vartheta, x)$ is analytic on $\pi_1(E)$. The interchange of the integration order is justified by (5.4) and the compactness of the contour γ. In the same manner one can establish the analyticity of the mapping $\mathbb{C} \ni x \mapsto F(\vartheta, x)$ for any fixed $\vartheta \in \pi_1(E)$. Summing up, we have shown that F is partially analytic on E and Hartogs' theorem now implies the assertion of the Lemma. □

5.5 Lemma: *For $c \in \mathbb{R}$ and $\sigma > 0$ all solutions of the second order differential equation*

$$f(x) - f''(x) = c\, x\, \mathrm{N}(0, \sigma^2, x)$$

are of the form

$$f_{k_1,k_2}(x) := k_1 e^x + k_2 e^{-x} + \frac{c\sigma^2}{4} e^{\frac{\sigma^2}{2}+x} \operatorname{erf}\left(\frac{x+\sigma^2}{\sqrt{2\sigma^2}}\right) + \frac{c\sigma^2}{4} e^{\frac{\sigma^2}{2}-x} \operatorname{erf}\left(\frac{x-\sigma^2}{\sqrt{2\sigma^2}}\right)$$

where $k_1, k_2 \in \mathbb{R}$ and $\operatorname{erf}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{y^2} \mathrm{d}y$. Moreover, we have $f_{k_1,k_2}(0) = k_1 + k_2$.

Proof. We observe that for $h(x) := c\, x\, \mathrm{N}(0, \sigma^2, x)$ the differential equation from above is equivalent to the first order inhomogeneous system

$$\begin{pmatrix} f \\ g \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f \\ g \end{pmatrix} + \begin{pmatrix} 0 \\ h \end{pmatrix}.$$

The assertion directly follows from the well-known uniqueness and existence theorems for first order ODEs, see for instance [WA, page 162]. □

5.2 Analytic dependence of zeros

The following factorization theorem from multivariate complex analysis gives a sufficient condition for the analytic dependence of zeros. It is a direct consequence of the Weierstrass preparation theorem. More details can be found in [CH, Chapter 1] .

5.6 Theorem: *For $n \geq 2$ let f be a analytic function on a domain $G = D' \times D \subset \mathbb{C}^n$ where $D' \subset \mathbb{C}^{n-1}$ is simply connected. Assume that the function f has for each $z' \in D'$ exactly m distinct zeros in the set D. Then there exist analytic functions $\alpha_1, ..., \alpha_m : D' \to D$, natural numbers $k_1, ..., k_m$ and an analytic function $\Phi : G \to \mathbb{C}$ that does not vanish on G such that*

$$f(z', z) = \prod_{l=1}^{m} (z - \alpha_l(z'))^{k_l} \, \Phi(z', z)$$

holds true for any $(z', z) \in G$.

The following version of the analytic implicit function theorem is well suited for our purposes. It can be obtained as a corollary of Theorem 5.6 by applying well-known ideas from the proof of Rouché's theorem, cf. [CO, page 125].

5.7 Theorem: *For $n \geq 2$ let f be an analytic function on a domain $G = D' \times D \subset \mathbb{C}^n$ where $D' \subset \mathbb{C}^{n-1}$ is simply connected. Assume that $f(z'_0, z_0) = 0$ and $(\partial_z f)(z'_0, z_0) \neq 0$ holds for some point $(z'_0, z_0) \in G$. Then we can find open neighborhoods $U(z'_0) \subset D'$ and $V(z_0) \subset D$ of z'_0 and z_0 as well as an analytic function $g : U(z'_0) \to V(z_0)$ such that*

$$f(z', z) = 0 \Leftrightarrow z = g(z')$$

holds for all $z' \in U(z'_0)$ and $z \in V(z_0)$.

Proof. Due to the assumption $(\partial_z f)(z'_0, z_0) \neq 0$, we can choose some $\varepsilon_1 > 0$ such that the ball $B_{\varepsilon_1}(z_0)$ is contained in D and $f(z'_0, z) \neq 0$ holds for any $z \in B_{\varepsilon_1}(z_0) \setminus \{z_0\}$. Moreover, we can pick constants $c, \varepsilon_2 > 0$ such that $B_{\varepsilon_2}(z'_0)$ is contained in D' and $|f(z', z)| > c$ holds for all $z' \in B_{\varepsilon_2}(z'_0)$ and any $z \in \mathbb{C}$ satisfying $|z - z_0| = \varepsilon_1$. By choice of ε_1 and the argument principle, we find that

$$\frac{1}{2\pi i} \oint_{|z-z_0|=\varepsilon_1} \frac{(\partial_z f)(z'_0, z)}{f(z'_0, z)} \, \mathrm{d}z = 1.$$

The triangle inequality for line integrals now yields

$$\begin{aligned}
&\sup_{|z'-z'_0|<\varepsilon_2/n} \left| 1 - \frac{1}{2\pi i} \oint_{|z-z_0|=\varepsilon_1} \frac{(\partial_z f)(z', z)}{f(z', z)} \, \mathrm{d}z \right| \\
&\quad = \frac{1}{2\pi} \sup_{|z'-z'_0|<\varepsilon_2/n} \left| \oint_{|z-z_0|=\varepsilon_1} \frac{(\partial_z f)(z'_0, z) f(z', z) - (\partial_z f)(z', z) f(z'_0, z)}{f(z'_0, z) f(z', z)} \, \mathrm{d}z \right| \\
&\quad \leq \alpha \sup_{\substack{|z'-z'_0|<\varepsilon_2/n \\ |z-z_0|=\varepsilon_1}} |(\partial_z f)(z'_0, z) f(z', z) - (\partial_z f)(z', z) f(z'_0, z)|
\end{aligned}$$

for any $n \in \mathbb{N}$. Here we denote by α some positive constant independent of n. Owing to the continuity of f and its derivatives, we conclude that the right-hand side of the latter inequality converges to 0 as n tends to infinity. Furthermore, the integral expression

$\frac{1}{2\pi i}\oint_{|z-z_0|=\varepsilon_1} \frac{(\partial_z f)(z',z)}{f(z',z)}\,\mathrm{d}z$ is integer valued and consequently we can find some $n_0 \in \mathbb{N}$ such that

$$\frac{1}{2\pi i}\oint_{|z-z_0|=\varepsilon_1} \frac{(\partial_z f)(z',z)}{f(z',z)}\,\mathrm{d}z = 1$$

holds true for any $z' \in B_{\varepsilon_2/n_0}(z_0')$. To put differently, for any $z' \in B_{\varepsilon_2/n_0}(z_0')$ the mapping $z \mapsto f(z',z)$ has exactly one zero within the set $B_{\varepsilon_1}(z_0)$. By virtue of Theorem 5.6, there exists an analytic function $g : B_{\varepsilon_2/n_0}(z_0') \to B_{\varepsilon_1}(z_0)$ such that $f(z',g(z')) = 0$. □

5.3 Basic notions from functional analysis

Let X denote some topological space. An extended real-valued function $f : X \to [-\infty,\infty]$ is called *lower semi-continuous* if for any $a \in \mathbb{R}$ the level set $\{x \in X \,|\, f(x) \le a\}$ is closed. Moreover, we say that f is *upper semi-continuous* if $-f$ is a lower semi-continuous function. It is easy to see that f is a lower semi-continuous function if and only if

$$f(x) \le \sup_{O\in\mathcal{U}(x)} \inf_{y\in O} f(y) \tag{5.8}$$

holds true for any $x \in X$. Here we denote by $\mathcal{U}(x)$ the collection of all open sets containing the point x. The latter statement can be rendered more precisely if X is a metric space. In this case we find that f is lower semi-continuous if and only if

$$f(x) \le \liminf_{n\to\infty} f(x_n)$$

for all $x \in X$ and all sequences $(x_n)_{n\in\mathbb{N}} \subset X$ converging to x. The following properties are easily derived from the definition of lower semi-continuity.

5.9 Lemma: *Suppose that f, g and $(f_i)_{i\in I}$ are lower semi-continuous functions where I denotes some non-empty index set. Then:*

1. *The functions $f+g$, $f \wedge g$ and $f \vee g$ are lower semi-continuous.*
2. *The mapping $x \mapsto \sup_{i\in I} f_i(x)$ is lower semi-continuous.*
3. *If f is bounded from below on some compact set K, then f attains its minimum on K.*
4. *The function f is lower semi-continuous if and only if its* epigraph

$$\operatorname{epi}(f) := \{(x,s) \in X \times \mathbb{R} \mid f(x) \le s\}$$

is closed with respect to the product topology on $X \times \mathbb{R}$.

Suppose that $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ and let V be a vector space over $\mathbb{K}$. Moreover, for an arbitrary index set I let $P := (p_i)_{i\in I} \subset \mathbb{R}_+^V$ denote some family of *semi-norms* on V. That is to say, for any index $i \in I$ and all $x, y \in V, \lambda \in \mathbb{K}$, we have $p_i(\lambda x) = |\lambda| p_i(x)$ and

$p_i(x+y) \leq p_i(x) + p_i(y)$. The initial topology on V induced by the semi-norms $(p_i)_{i\in I}$ is given by

$$\mathcal{T}_P := \Big\{ O \subset V \,\Big|\, \forall\, x_0 \in O \;\exists\, F \subset I \text{ finite } \exists\, \varepsilon \in (0,\infty)^F : B^F_\varepsilon(x_0) \subset O \Big\} \tag{5.10}$$

where $B^F_\varepsilon(x_0) := \cap_{i\in F}\, \{x \in V \mid p_i(x - x_0) < \varepsilon_i\}$. The latter topology is the coarsest topology on the vector space V such that for all $i \in I, x_0 \in X$ the mappings $x \mapsto p_i(x - x_0)$ are continuous. We call $\mathcal{T}_P$ the *locally convex topology* on V generated by the semi-norm family P and $(V, \mathcal{T}_P)$ a *locally convex (topological vector) space.* It is easy to verify that $(V, \mathcal{T}_P)$ is indeed a topological vector space in the classical sense, i.e. vector addition and scalar multiplication constitute continuous operations.

There exists an equivalent approach towards locally convex spaces which is rather geometrically motivated, cf. [RD, Chapter 1]. An introduction via semi-norms similar to (5.10) can be found in [CO2, Chapter 4]. Let us remark that the definition of $\mathcal{T}_P$ from above does not ensure that locally convex topologies are always Hausdorff. Neither do locally convex topologies always need to be metrizable. Nevertheless, the following characterizations are well-known.

5.11 Lemma: *Let V and P be defined as above.*

1. *The topology $\mathcal{T}_P$ is Hausdorff if and only if the family of semi-norms is point separating, i.e. for all $x \in V$ exists an index $i \in I$ such that $p_i(x) \neq 0$.*
2. *The topology $\mathcal{T}_P$ is metrizable if and only if $\mathcal{T}_P$ is generated by a point separating, countable family $(\tilde{p}_n)_{n\in\mathbb{N}}$ of semi-norms. In this case the compatible metric is given by*
$$d(x,y) := \sum_{n=0}^{\infty} 2^{-n} \frac{\tilde{p}_n(x-y)}{1 + \tilde{p}_n(x-y)}.$$

For details about metrizable and normable locally convex spaces see Section 2 of [CO2, Chapter 4]. The following facts concerning continuity and convergence in locally convex spaces are essentially a reformulation of Proposition 7.7 and Proposition 7.8 from [TR].

5.12 Lemma: *Suppose that V and W are $\mathbb{K}$-vector spaces endowed with the locally convex topologies generated by some semi-norm families $P := (p_i)_{i\in I} \subset \mathbb{R}^V_+$ and $Q := (q_j)_{j\in J} \subset \mathbb{R}^W_+$. Then:*

1. *A net $x_\alpha \subset V$ converges to $x \in V$ with respect to $\mathcal{T}_P$ if and only if for any index $i \in I$ the net $p_i(x_\alpha - x)$ converges to 0 with respect to the Euclidean topology on $\mathbb{R}$.*
2. *Let $T : V \to W$ be a linear mapping. The following statements are equivalent:*
   a) *T is continuous.*
   b) *T is continuous at 0.*
   c) *For any net $x_\alpha \subset V$ such that $p_i(x_\alpha) \to 0$ for all $i \in I$, we have $q_j(Tx_\alpha) \to 0$ for all $j \in J$.*

d) *For any index $j \in J$ exists a finite set $F \subset I$ and some constant $c \geq 0$ such that $q_j(Tx) \leq c \max_{i\in F} p_i(x)$ holds true for any $x \in V$.*

The lemma above relies upon the notion of net convergence. A brief revision of the required theory concerning nets in topological spaces can be found in the second paragraph of [CO2, Appendix A]. As a corollary of Lemma 5.12 we obtain that a linear functional $x^* : V \to \mathbb{K}$ is $\mathcal{T}_P$-continuous if and only if there exists a finite set $F \subset I$ and a constant $c \geq 0$ such that $|x^*(x)| \leq c \max_{i\in F} p_i(x)$ for all $x \in V$. The vector space of all $\mathcal{T}_P$-continuous linear functionals on V is called *(continuous) dual space* and denoted by $(V, \mathcal{T}_P)^*$. If there is no possibility of misinterpretation, we occasionally refrain from explicitly indicating the topology and simply write V^*. Locally convex spaces are in some sense a natural extension of normed spaces. Indeed, for any normed vector space $(V, \|\cdot\|)$ and $P := \{\|\cdot\|\}$, it is easy to see that $\mathcal{T}_P$ coincides with the norm topology on V. Locally convex topologies are stable with respect to certain topological operations.

5.13 Lemma:

1. *Suppose that $(V_j, \mathcal{T}_{P_j})_{j\in J}$ is a family of locally convex space over $\mathbb{K}$ where J denotes some index set. The Cartesian product $\times_{j\in J} V_j$ is a locally convex space with respect to the product topology $\bigotimes_{j\in J} \mathcal{T}_{P_j}$. More precisely, we have*

$$\bigotimes_{j\in J} \mathcal{T}_{P_j} = \mathcal{T}_Q$$

for the family of semi-norms $Q := \bigcup_{j\in J}\{\, p \circ \pi_j \mid p \in P_j\}$ where $\pi_i : \times_{j\in J} V_j \to V_i$ denotes the canonical projection onto V_i.

2. *Any linear subspace U of some locally convex space $(V, \mathcal{T}_P)$ constitutes a locally convex space with respect to the subspace topology $\mathcal{T}_P \cap U$. More precisely, we have*

$$\mathcal{T}_P \cap U = \mathcal{T}_{P|_U}$$

for the family of semi-norms $P|_U := \{\, p : U \to \mathbb{R}_+ \mid p \in P\}$.

Locally convex topologies possess all required geometric properties such that Hahn-Banach type theorems hold true. We summarize some results from Section 3 of [CO2, Chapter 4] and [TR, Chapter 18].

5.14 Lemma: *Let $(V, \mathcal{T}_P)$ denote some locally convex space over $\mathbb{K}$.*

1. *Let U be a linear subspace of V and suppose that $u^* \in (U, \mathcal{T}_{P|_U})^*$. There exists a continuous, linear functional $x^* \in (V, \mathcal{T}_P)^*$ such that $x^*|_U = u^*$.*

2. *Suppose that $C_1, C_2 \subset V$ are disjoint, convex sets and C_1 is open. Then there exists a continuous, linear functional $x^* \in (V, \mathcal{T}_P)^*$ such that $\operatorname{Re} x^*(x_1) < \operatorname{Re} x^*(x_2)$ for all $x_1 \in C_1$ and $x_2 \in C_2$.*

3. *Suppose that $C \subset V$ is a closed, convex set and that $x \notin C$. We can find some $x^* \in (V, \mathcal{T}_P)^*$ and $\varepsilon > 0$ such that $\operatorname{Re} x^*(x) + \varepsilon \leq \operatorname{Re} x^*(x_1)$ for all $x_1 \in C$.*

4. *If the topology $\mathcal{T}_P$ is Hausdorff, then for any choice of $x, y \in V, x \neq y$ there exists some $x^* \in (V, \mathcal{T}_P)^*$ such that $x^*(x) \neq x^*(y)$.*

Suppose that $\langle V, W\rangle$ is an *algebraic pairing* over the field $\mathbb{K}$, i.e. V and W are $\mathbb{K}$-vector spaces and $\langle\cdot,\cdot\rangle : V \times W \to \mathbb{K}$ is a bilinear mapping. It is easy to see that $P := \{x \mapsto |\langle x, y\rangle| \mid y \in W\}$ defines a family of semi-norms on V. The locally convex topology on V generated by P is called the *weak topology* on V induced by W and denoted by $\sigma(V, W)$. This concept is commutative in the sense that V induces in the same manner a topology on the vector space W, i.e. the weak topology $\sigma(W, V)$. The following Lemma is folklore.

5.15 Lemma: *Let $\langle V, W\rangle$ be an algebraic pairing over $\mathbb{K}$, then:*

1. *If W separates the points of the vector space V, i.e. for any $x \in V \setminus \{0\}$ there exists some $y \in W$ such that $\langle x, y\rangle \neq 0$, then the topology $\sigma(V, W)$ is Hausdorff.*

2. *A linear functional $x^* : V \to \mathbb{K}$ is $\sigma(V, W)$-continuous if and only if there exists a vector $y \in W$ such that $x^*(x) = \langle x, y\rangle$. Consequently, we have*
$$(V, \sigma(V, W))^* = W.$$

3. *A net $x_\alpha \subset V$ converges to $x \in V$ with respect to $\sigma(V, W)$ if and only if $\langle x_\alpha, y\rangle \to \langle x, y\rangle$ for any $y \in W$.*

The roles of the spaces V and W can be interchanged in all the statements above.

Clearly, any locally convex space $(V, \mathcal{T}_P)$ is paired with its continuous dual V^* via the bilinear mapping $\langle x, x^*\rangle := x^*(x)$ where $x \in V$ and $x^* \in V^*$. The points of the dual space V^* are always separated by V and consequently Lemma 5.15 warrants that $\sigma(V^*, V)$ is Hausdorff. Besides, if $(V, \mathcal{T}_P)$ is a Hausdorff space, we can conclude by Lemma 5.14 that the points of V are separated by V^* and therefore $\sigma(V, V^*)$ is Hausdorff as well.

Two topologies which play a distinguished role in probability theory are generated by algebraic pairings. Let $\mathcal{M}(\mathbb{R}^n)$ be the vector space of regular Borel measures on $\mathbb{R}^n$ with finite total variation. Moreover, denoted by $C_0(\mathbb{R}^n)$ and $C_b(\mathbb{R}^n)$ the continuous functions vanishing at infinity and the bounded continuous functions on $\mathbb{R}^n$, respectively. It is easy to see that
$$\langle f, \mu\rangle \mapsto \int_{\mathbb{R}^n} f \mathrm{d}\mu$$
defines a point separating, bilinear mapping on the Cartesian products $C_0 \times \mathcal{M}$ and $C_b \times \mathcal{M}$. Lemma 5.15 therefore warrants both topologies are Hausdorff. The locally convex topologies $\sigma(\mathcal{M}, C_b)$ and $\sigma(\mathcal{M}, C_0)$ correspond to the *measure theoretic weak and vague topology* on $\mathcal{M}(\mathbb{R}^n)$. Let us remark that the weak topology is metrizable on the set of probability measures $\mathcal{P}(\mathbb{R}^n)$, cf. [KL, Remark 13.14]. Furthermore, the vague topology is metrizable on the total variation unit ball $B_{\mathcal{M}(\mathbb{R}^n)}$. Indeed, choose a sequence $f_1, f_2, \ldots$ which is dense in $C_0(\mathbb{R}^n)$ and define $d(\mu, \nu) := \sum_{n\in\mathbb{N}} 2^{-n}|\langle f_n, \mu - \nu\rangle|$ for any

$\mu, \nu \in B_{\mathcal{M}(\mathbb{R}^n)}$. The latter metric generates the subspace topology on $B_{\mathcal{M}(\mathbb{R}^n)}$ induced by $\sigma(\mathcal{M}, C_0)$. In virtue of Lemma 5.15, we have that $(C_0(\mathbb{R}^n), \sigma(C_0, \mathcal{M}))' = \mathcal{M}(\mathbb{R}^n)$. From the Riesz representation theorem we know that the topology of uniform convergence is compatible with this duality. To be more specific, the continuous dual of the Banach space $(C_0(\mathbb{R}^n), \|\cdot\|_\infty)$ is isometrically isomorph to the regular Borel measures $\mathcal{M}(\mathbb{R}^n)$ equipped with the total variation norm, cf. [RD2, Theorem 6.19]. Clearly, the topology $\sigma(C_0, \mathcal{M})$ is coarser than the topology of uniform convergence.

The concept of *dual operators* generalizes in a natural way to paired spaces. Suppose that $\langle V_1, W_1\rangle_1$ and $\langle V_2, W_2\rangle_2$ are algebraic pairings and let $T : V_1 \to V_2$ denote some $\sigma(V_1, W_1)$-$\sigma(V_2, W_2)$-continuous, linear mapping. By virtue of Lemma 5.15 and the Hahn-Banach type results from Lemma 5.14, it is easy to show that there exists a uniquely determined, $\sigma(W_2, V_2)$-$\sigma(W_1, V_1)$-continuous, linear operator $T^* : W_2 \to W_1$ such that

$$\langle Tx, y\rangle_2 = \langle x, T^* y\rangle_1 \tag{5.16}$$

holds true for arbitrary $x \in V_1$ and $y \in W_2$. We call T^* the *dual* or *adjoint operator* associated to T. Interestingly, we have the following Hellinger-Toeplitz type converse result.

5.17 Lemma: *Let $\langle V_1, W_1\rangle_1$ and $\langle V_2, W_2\rangle_2$ be algebraic pairings. Suppose that $T : V_1 \to V_2$ and $T^* : W_2 \to W_1$ are linear mappings satisfying Equation* (5.16) *for all $x \in V_1$ and $y \in W_2$. Then T is $\sigma(V_1, W_1)$-$\sigma(V_2, W_2)$-continuous and T^* is $\sigma(W_2, V_2)$-$\sigma(W_1, V_1)$-continuous.*

Proof. Let x_α be a net that converges to $x \in V_1$ with respect to $\sigma(V_1, W_1)$. The third assertion of Lemma 5.15 and Equation (5.16) imply that $\lim\langle Tx_\alpha, y\rangle = \lim\langle x_\alpha, T^* y\rangle = \langle x, T^* y\rangle = \langle Tx, y\rangle$ holds for any $y \in W_2$. Again by Lemma 5.15 we conclude that the net Tx_α converges to Tx in the $\sigma(V_2, W_2)$ topology. This proves the desired continuity of the mapping T. The continuity of the operator T^* can be easily established along the same lines. □

An important feature of weak topologies is that they are in a certain way rich of compact sets. The following theorem might be considered as one of the fundamental results from functional analysis.

5.18 Theorem (Alaoglu-Bourbaki)**:** *Suppose that V is a Hausdorff locally convex space and let $U \subset V$ be a neighborhood of* 0. *The polar set*

$$U^\circ := \{x^* \in V^* \mid \operatorname{Re} x^*(x) \leq 1 \; \forall\, x \in U\}$$

is $\sigma(V^*, V)$-compact.

A proof of the latter theorem can be found in [MV, Theorem 23.5]. In the case that $(V, \|\cdot\|)$ is a Banach space, the original version of Alaoglu's theorem is easily recovered from Theorem 5.18. Indeed, choosing $U = B_V := \{x \in V \mid \|x\| \leq 1\}$ yields that the dual unit ball $U^\circ = B_{V^*} := \{x^* \in V^* \mid \|x^*\|_* \leq 1\}$ is *weak*-compact*, i.e. compact with respect to the topology $\sigma(V^*, V)$. Here we denote by $\|x^*\|_* := \sup_{\|x\|=1} x^*(x)$ the dual norm on the space V^*. As an example we apply Theorem 5.18 to the pairing $\langle C_0(\mathbb{R}^n), \mathcal{M}(\mathbb{R}^n)\rangle$ from above and obtain that the total variation unit ball $B_{\mathcal{M}(\mathbb{R}^n)}$ is vaguely compact.

5.4 Convex conjugation and the Fenchel-Moreau theorem

This section aims at presenting some basic principles from convex duality theory in a nutshell. We rely on the analytic notions introduced in Section 5.3. An extended real-valued function $f : V \to [-\infty, \infty]$ defined on some real vector space V is called *convex* if its epigraph $\mathrm{epi}(f) = \{(x, s) \in V \times \mathbb{R} \mid f(x) \leq s\}$ is a convex subset of $V \times \mathbb{R}$. In the case that f only assumes values in $\mathbb{R}$, the latter definition is equivalent to the classical notion of convexity, i.e. $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ for all $\lambda \in [0, 1]$ and $x, y \in V$. Convex functions possess a stability property that we have already encountered in the context of lower semi-continuous functions, cf. Lemma 5.9.

5.19 Lemma: *Suppose that $(f_i)_{i\in I}$ is a family of extended real-valued, convex functions on V. The point-wise supremum $x \mapsto \sup_{i\in I} f_i(x)$ defines a convex mapping.*

Now suppose that $(V, \mathcal{T}_P)$ is a locally convex space over $\mathbb{R}$ and denote by $\langle x, x^*\rangle := x^*(x)$ the canonical pairing on $V \times V^*$. For any $x^* \in V^*$ and $r \in \mathbb{R}$, we call the affine mapping $\phi(x) := r + \langle x, x^*\rangle$ a *supporting functional* of some set $M \subset V$ if $\sup_{x\in M} \phi(x) = 0$. In other words, the set M is contained in the closed half-space $H_\phi := \{x \in V \mid \phi(x) \leq 0\}$ and there is no margin between M and the *supporting hyperplane* $\{x \in V \mid \phi(x) = 0\}$. In virtue of the Hahn-Banach separation theorem, we can uniquely encode any non-empty, closed, convex set $C \subset V$ in terms of its supporting functionals. Indeed, the third assertion of Lemma 5.14 yields that a point $x_0 \in V$ is contained in the complement of C if and only if there exists a continuous, linear functional $x^* \in V^*$ such that $\langle x_0, x^*\rangle > \langle x, x^*\rangle$ holds true for all $x \in C$. Clearly, for any $x^* \in V^*$ with $\sup_{x\in C}\langle x, x^*\rangle < \infty$ the mapping

$$x \mapsto \langle x, x^*\rangle - \sup_{x'\in C}\langle x', x^*\rangle$$

constitutes a supporting functional of the set C. We conclude that $x_0 \in V$ is contained in C^{c} if and only if there exists a supporting functional ϕ of C such that $\phi(x_0) > 0$ and therefore we have

$$C^{\mathrm{c}} = \bigcup_{\phi \text{ supp. func.}} H_\phi^{\mathrm{c}} = \bigcup_{x^*\in V^*} \{\, x \in V \mid \langle x, x^*\rangle > \sup_{x'\in C}\langle x', x^*\rangle \,\}.$$

Taking the complement yields that the set C can be recovered by intersecting the closed half-spaces H_ϕ associated to the supporting functionals of C, i.e.

$$C = \bigcap_{\phi \text{ supp. func.}} H_\phi = \bigcap_{x^*\in V^*} \{\, x \in V \mid \langle x, x^*\rangle \leq \sup_{x'\in C}\langle x', x^*\rangle \,\}. \tag{5.20}$$

The mapping $\mathcal{I}_C^* : V^* \to (-\infty, \infty]$ defined by

$$\mathcal{I}_C^*(x^*) := \sup_{x \in C} \langle x, x^* \rangle = \sup_{x \in V} \{ \langle x, x^* \rangle - \mathcal{I}_C(x) \} \tag{5.21}$$

has the following simple geometric interpretation: For any slope $x^* \in V^*$ with $\mathcal{I}_C^*(x^*) < \infty$ and any intercept $r \in \mathbb{R}$ consider the affine mapping $x \mapsto \langle x, x^* \rangle + r$. The unique intercept such that the latter mapping constitutes a supporting functional of the set C is given by $r := -\mathcal{I}_C^*(x^*)$. In this sense, all information concerning the supporting functionals of C is encoded within the mapping $\mathcal{I}_C^*$. In virtue of Equation (5.20), it is not surprising that two closed, convex sets $C_1, C_2 \subset V$ coincide if and only if $\mathcal{I}_{C_1}^*(x^*) = \mathcal{I}_{C_2}^*(x^*)$ for all $x^* \in V^*$.

The observation that any closed, convex set C is uniquely determined by the associated *conjugate indicator* $\mathcal{I}_C^*$ from (5.21) can be considered as a cornerstone of the convex duality theory originating from the work of Adrien-Marie Legendre, Hermann Minkowski and Werner Fenchel. From Lemma 5.12 we know that a function $f : V \mapsto (-\infty, \infty]$ is lower semi-continuous if and only if its epigraph $\mathrm{epi}(f) = \{(x, s) \in V \times \mathbb{R} \mid f(x) \leq s\}$ is closed with respect to the product topology on $V \times \mathbb{R}$. Recall that $V \times \mathbb{R}$ endowed with the product topology $\mathcal{T}_P \otimes \mathcal{T}_\mathbb{R}$ constitutes a locally convex space, see Lemma 5.13. Moreover, the reader easily verifies that f is a convex function if and only if $\mathrm{epi}(f)$ is a convex subset of $V \times \mathbb{R}$. Now suppose that $f : V \mapsto (-\infty, \infty]$ is a lower semi-continuous, convex function. The discussion from above suggests that the epigraph of f, and therefore the function f itself, is uniquely characterized by the associated supporting functionals. The Fenchel-Moreau biconjugate theorem below renders this statement more precisely.

In order to formulate the biconjugate theorem, we require some fundamental notions from convex duality theory. For a detailed exposition we refer the reader to [RO, Section 3]. Suppose that $(V, \mathcal{T}_P)$ is a locally convex space over $\mathbb{R}$ and denote by $f : V \to [-\infty, \infty]$ some extended real-valued mapping. The *conjugate* $f^* : V^* \to [-\infty, \infty]$ of the function f is defined by

$$f^*(x^*) := \sup_{x \in V} \{ \langle x, x^* \rangle - f(x) \} \tag{5.22}$$

and the operation $f \mapsto f^*$ is called the *Fenchel transform*. The symmetry inherent in the algebraic pairing $\langle V, V^* \rangle$ allows us to iterate the conjugation process. A function of particular interest is the *biconjugate* $f^{**} : V \to [-\infty, \infty]$ associated to f given by

$$f^{**}(x) := (f^*)^*(x) = \sup_{x^* \in V^*} \{ \langle x, x^* \rangle - f^*(x^*) \}.$$

Moreover, the mappings $\mathrm{lsc}(f), \mathrm{co}(f) : V \to [-\infty, \infty]$ defined by

$$\mathrm{lsc}(f)(x) := \sup\{ h(x) \mid h \text{ lower semi-continuous and } h \leq f \}, \tag{5.23}$$
$$\mathrm{co}(f)(x) := \sup\{ h(x) \mid h \text{ convex and } h \leq f \} \tag{5.24}$$

are called the *lower semi-continuous envelope* and the *convex envelope* of the function f, respectively. Following [RO, Section 3], we define by

$$\mathrm{cl}(f) := \begin{cases} \mathrm{lsc}(f) & \text{if } \mathrm{lsc}(f)(x) > -\infty \text{ for all } x \in V, \\ -\infty & \text{otherwise} \end{cases}$$

the *closure* of the function f. Moreover, we say that f is a *closed* function if $f = \mathrm{cl}(f)$. Let us collect some simple properties of the latter objects.

5.25 Lemma:

1. *For any $x \in V$ and $x^* \in V^*$, we have $\langle x, x^* \rangle \leq f(x) + f^*(x^*)$ and $f^{**}(x) \leq f(x)$.*

2. *The conjugate and biconjugate of f are convex and lower semi-continuous.*

3. *The lower semi-continuous envelope of f is the largest lower semi-continuous minorant of f. In other words, the function $\mathrm{lsc}(f)$ is lower semi-continuous, $\mathrm{lsc}(f) \leq f$ and $h \leq \mathrm{lsc}(f)$ for any other lower semi-continuous function h satisfying $h \leq f$. For any $x \in V$, we have*
$$\mathrm{lsc}(f)(x) = \sup_{O \in \mathcal{U}(x)} \inf_{y \in O} f(y) \tag{5.26}$$
where $\mathcal{U}(x)$ denotes the collection of all open sets containing x. Moreover, the epigraph of $\mathrm{lsc}(f)$ coincides with the closure of $\mathrm{epi}(f)$ with respect to the product topology $\mathcal{T}_P \otimes \mathcal{T}_{\mathbb{R}}$ on $V \times \mathbb{R}$, i.e.
$$\mathrm{epi}(\mathrm{lsc}(f)) = \mathrm{cl}(\mathrm{epi}(f)). \tag{5.27}$$

4. *The convex envelope of f is the largest convex minorant of f. In other words, the function $\mathrm{co}(f)$ is convex, $\mathrm{co}(f) \leq f$ and $h \leq \mathrm{co}(f)$ for any other convex function h satisfying $h \leq f$. Moreover, the epigraph of $\mathrm{co}(f)$ coincides with the convex hull of the set $\mathrm{epi}(f)$, i.e.*
$$\mathrm{epi}(\mathrm{co}(f)) = \mathrm{co}(\mathrm{epi}(f)). \tag{5.28}$$

Proof.

1. We observe that the inequality $f^*(x^*) \geq \langle x, x^* \rangle - f(x)$ holds by definition of the conjugate. Moreover, we have $f^{**}(x) = \sup_{x^* \in V^*} \inf_{y \in V} \{\langle x - y, x^* \rangle + f(y)\} \leq f(x)$ for any $x \in V$.

2. This assertion follows directly from Lemma 5.9 and Lemma 5.19 by noting that the affine mappings occurring in the definition of f^* and f^{**} are convex and continuous.

3. The lower semi-continuity of the function $\mathrm{lsc}(f)$ as defined in (5.23) follows directly from the second assertion of Lemma 5.9. Moreover, Equation (5.23) directly implies that $h \leq \mathrm{lsc}(f) \leq f$ holds true for any lower semi-continuous function h such that $h \leq f$. In other words, the function $\mathrm{lsc}(f)$ corresponds indeed to the

largest lower semi-continuous minorant of f. In order to verify the representation formula (5.26), we observe that $g(x) := \sup_{O\in\mathcal{U}(x)} \inf_{y\in O} f(y)$ constitutes a lower semi-continuous minorant of f. For any other lower semi-continuous function h with $h \leq f$, we obtain from (5.8) that

$$h(x) \leq \sup_{O\in\mathcal{U}(x)} \inf_{y\in O} h(y) \leq \sup_{O\in\mathcal{U}(x)} \inf_{y\in O} f(y) = g(x).$$

This yields that g is the largest lower semi-continuous minorant of f and therefore $g = \operatorname{lsc}(f)$ holds true. It remains to verify Equation (5.27). The inclusion $\operatorname{cl}\operatorname{epi}(f) \subset \operatorname{epi}(\operatorname{lsc}(f))$ is obvious, since the epigraph of $\operatorname{lsc}(f)$ is a closed superset of $\operatorname{epi}(f)$, cf. Lemma 5.9. In order to prove the reverse inclusion, choose any $(x, s) \in \operatorname{epi}(\operatorname{lsc}(f))$. Equation (5.26) implies that for any $O \in \mathcal{U}(x)$ and any $\varepsilon > 0$ exists some $y \in O$ such that $f(y) \leq s + \varepsilon$. To put it another way, for any $\mathcal{T}_P \otimes \mathcal{T}_\mathbb{R}$-open neighborhood $O \times (s-\varepsilon, s+\varepsilon) \in \mathcal{U}(x, s)$ there exists some point $(y, \tilde{s}) \in O \times (s-\varepsilon, s+\varepsilon)$ such that $(y, \tilde{s}) \in \operatorname{epi}(f)$. This yields that the point (x, s) is contained in the closure of the set $\operatorname{epi}(f)$.

4. This assertion is certainly true for any function f which assumes the value $-\infty$ at some point. Indeed, in this case we have $\operatorname{co}(f) = -\infty$ and therefore $\operatorname{epi}(\operatorname{co}(f)) = \operatorname{co}(\operatorname{epi}(f)) = V \times \mathbb{R}$. Now suppose that $f : V \to (-\infty, \infty]$. In virtue of Equation (5.24) and Lemma 5.19, it is apparent that $\operatorname{co}(f)$ corresponds to the largest convex minorant of f and that $\operatorname{epi}(\operatorname{co}(f))$ is a convex superset of $\operatorname{epi}(f)$. In order to obtain (5.28), the reader only needs to observe that

$$\operatorname{co}(f)(x) = \inf\{ s \in (-\infty, \infty] \mid (x, s) \in \operatorname{co}(\operatorname{epi}(f)) \}$$

holds true for any $x \in V$.

□

The following theorem is an adaption of [RO, Theorem 5]. Variations of the latter have become known as the *Fenchel-Moreau theorem* or the *biconjugate theorem*.

5.29 Theorem (Fenchel-Moreau)**:** *Suppose that $f : V \to [-\infty, \infty]$ is an extended real-valued mapping, then:*

1. *The conjugate f^* constitutes a closed, convex function on the dual space V^* and we have $f^{**} = \operatorname{cl}(\operatorname{co}(f))$.*
2. *The Fenchel transform induces a one-to-one correspondence between the closed, convex functions on V and the closed, convex functions on V^*.*

The latter theorem generalizes in some sense our discussion from the beginning of this section. Indeed, assume that f and g are closed, convex functions such that $f^* = g^*$. Owing to Theorem 5.29, we have $f = \operatorname{cl}(\operatorname{co}(f)) = f^{**} = g^{**} = \operatorname{cl}(\operatorname{co}(g)) = g$.

Bibliography

[AL] C. D. Aliprantis: Infinite Dimensional Analysis - A Hitchhiker's Guide, 3rd Edition, Springer, 2006

[ALO] L. Andersen, M. Lake, D. Offengenden: High Performance American Option Pricing, ssrn.com, 2015

[BG] M. Broadie, P. Glasserman: Pricing American-style securities using simulation, Journal of Economic Dynamics and Control 21, 1997

[BO] P. Boyle: Options - a Monte Carlo approach, Journal of Financial Economics 4, p. 323-338, 1977

[BRS] M. J. Brennan, E.S. Schwartz: The Valuation of American Put Options, The Journal of Finance, Volume 32, Issue 2, 1977

[BS] F. Black, M. Scholes: The Pricing of Options and Corporate Liabilities, The Journal of Political Economy, Vol. 81, No. 3, 1973

[CH] E. M. Chirka: Complex Analytic Sets, Kluwer Academic Publishers, 1989

[CJM] P. Carr, R. Jarrow, R. Myneni, Alternative characterization of American put Options, Mathematical Finance, Vol. 2, Issue 2, p. 87-106, 1992

[CLP] E. Clement, D. Lamberton, P. Protter: An analysis of the Longstaff-Schwartz algorithm for American option pricing, Finance and Stochastics, Vol. 6, p. 449-471, 2002

[CO] J. B. Conway: Functions of One Complex Variable I, Springer, 1978

[CO2] J. B. Conway: A Course in Functional Analysis, 2nd edition, Springer, 1990

[CR] S. Christensen: A Method for pricing American options using semi-infinite linear programming, Mathematical Finance, Vol. 24, Issue 1, 2014

[CRR] J. Cox, S. A. Ross, M. Rubinstein: Option pricing: a simplified approach, Journal of Financial Economics 7, p. 229-264, 1979

[DK] M. H. A. Davis and I. Karatzas: A deterministic approach to optimal stopping with applications, Probability Statistics and Optimization: a Tribute to Peter Whittle ed. F. P. Kelly, Wiley, New York and Chichester, p. 455-466, 1994

[EN] K. J. Engel, R. Nagel: One-Parameter Semigroups for Linear Evolution Equations, Springer, 2000

[GL] P. Glasserman: Monte Carlo Methods in Financial Engineering, Springer, 2003

[HK] R. Hettich, K. O. Kortanek: Semi-Infinite Programming - Theory, Methods and Applications, SIAM Review, Vol. 35, No. 3, 1993

[IW] S. Ito, S. Wu, T. J. Shiu, K. L. Teo: A numerical approach to infinite dimensional linear programming in L_1-spaces, Industrial and Management Optimization, Vol. 6, No. 1, 2010

[JLL] P. Jaillet, D. Lamberton, B. Lapeyre: Variational inequalities and the pricing of American options, Acta Appl. Math. 21, No. 3, p. 263-289, 1990

[JM1] B. Jourdain, C. Martini: American prices embedded in European prices, Annales de l'Institut Henri Poincaré, Non Linear Analysis, Vol. 18, No. 1, p. 1-17, 2001

[JM2] B. Jourdain, C. Martini: Approximation of American Put Prices by European Prices via an Embedding Method, Annals of Applied Probability, Vol. 12, No. 1, p. 196-223, 2002

[KB] O. Kallenberg: Foundations of Modern Probability, 2nd Edition, Springer, 2002

[KH] M. Haugh, L. Kogan: Pricing American options - a duality approach, Operations Research, Vol. 52, No. 2 p. 258-270, 2004

[KL] A. Klenke: Probability Theory - A Comprehensive Course, Springer, 2008

[KR] S. G. Krantz: Function Theory of Several Complex Variables, AMS Chelsea Publishing, 1992

[KU] H. J. Kushner: Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, 1977

[LA] S. Lang: Real and Functional Analysis, 3rd Edition, Springer, 1993

[LB] D. Lamberton: Optimal stopping and American options, Université Paris-Est, Laboratoire d'analyse et de mathématiques appliquées, lecture notes, 2009

[LL] D. Lamberton, B. Lapeyre, An introduction to stochastic calculus applied to finance, Chapman and Hall, 1996

[LS] F. A. Longstaff, E. S. Schwartz: Valuating American Options by Simulation - A Simple Least Square Approach, Review of Financial Studies, Vol. 14, No. 1, p.113-147, 2001

[LU] H. R. Lerche, M. Urusov: Optimal stopping via measure transformation - the Beibel-Lerche approach, Stochastics, Vol. 79, p. 275-291, 2007

[LW] H. C. Lai, S. Wu: Extremal points and optimal solutions for general capacity problems, Mathematical Programming, Vol. 54, p.87-113, 1992

[LZ] J. Lelong, A. Zanette: Tree methods, Encyclopedia of Quantitative Finance, p. 1829 - 1834, Wiley, 2010

[ME] R. C. Merton: The theory of rational option pricing, Bell Journal of Economics and Management Science, Vol. 4, No. 1, p. 141-183, 1973

[MV] R. Meise, D. Vogt: Introduction to Functional Analysis, Oxford University Press, 1997

[MX] C. Le Merdy, Quanhua Xu: Maximal theorems and square functions for analytic operators on L_p-spaces, Journal of the London Mathematical Society, Vol. 86, No. 2, p.343-365, 2012

[NP] C. P. Niculescu, L.-E. Persson: Convex functions and their applications, Springer, 2004

[PS] G. Peskir, A. N. Shiryaev: Optimal Stopping and Free-Boundary Problems, Lectures in Mathematics, ETH Zürich, Birkhäuser, 2006

[RD] W. Rudin: Functional Analysis, 2nd edition, McGraw-Hill, 1991

[RD2] W. Rudin: Real and Complex Analysis, 3rd edition, McGraw-Hill, 1986

[RG] L. C. G. Rogers: Monte Carlo valuation of American options, Mathematical Finance, Vol. 12, Issue 3, p. 271-286, 2002

[RO] R. T. Rockafellar: Conjugate Duality and Optimization, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 16, 1974

[RO2] R. T. Rockafellar: Convex Analysis, Princeton University Press, 1970

[RW] L. C. G. Rogers, D. Williams: Diffusions, Markov Processes and Martingales: Volume 1 - Foundations, 2nd Edition, Cambridge University Press, 1994

[SD] R. Seydel: Tools for Computational Finance, fourth edition, Springer, 2009

[SH] S. E. Shreve: Stochastic Calculus for Finance, Volume 1: The Binomial Asset Pricing Model, Springer, 2005

[SI] A. N. Shiryaev: Optimal Stopping Rules, Springer, 1978

[ST] E.M. Stein: Topics in Harmonic Analysis Related to the Littlewood-Paley Theory, Princeton University Press, 1970

[TR] F. Treves: Topological Vector Spaces, Distributions and Kernels, Academic Press, 1967

[TV] J. Toivanen: Finite difference methods for early exercise options, Encyclopedia of Quantitative Finance, p. 695-704, Wiley, 2010

[WA] W. Walter: Ordinary Differential Equations, Springer, 1998

[WB] G. Wahba: Spline Models for Observational Data, Springer, CBMS-NSF Regional Conference Series in Applied Mathematics, 1990

List of Figures

List of Tables

Erklärung

Hiermit erkläre ich, dass ich die vorliegende Dissertation - abgesehen von der Beratung durch meinen Betreuer Herrn Prof. Dr. Jan Kallsen - nach Inhalt und Form eigenständig angefertigt habe. Dabei habe ich die Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft eingehalten. Die Arbeit hat weder ganz noch zum Teil einer anderen Stelle im Rahmen eines Prüfungsverfahrens vorgelegen und ist weder ganz noch zum Teil veröffentlicht oder zur Veröffentlichung eingereicht worden.

Matthias Lenga
Kiel, 30. März 2017