ASYMPTOTICS AND DUALITY FOR THE DAVIS AND NORMAN PROBLEM

STEFAN GERHOLD, JOHANNES MUHLE-KARBE, AND WALTER SCHACHERMAYER

Abstract. We revisit the problem of maximizing expected logarithmic utility from consumption over an infinite horizon in the Black-Scholes model with proportional transaction costs, as studied in the seminal paper of Davis and Norman [*Math. Operation Research*, 15, 1990]. Similarly to Kallsen and Muhle-Karbe [*Ann. Appl. Probab.*, 20, 2010], we tackle this problem by determining a *shadow price*, that is, a frictionless price process with values in the bid-ask spread which leads to the same optimization problem. However, we use a different parametrization, which facilitates computation and verification. Moreover, for small transaction costs, we determine fractional Taylor expansions of arbitrary order for the boundaries of the no-trade region and the value function. This extends work of Janeček and Shreve [*Finance Stoch.*, 8, 2004], who determined the leading terms of these power series.

1. Introduction

It is a classical problem of financial theory to maximize expected utility from consumption (cf., e.g., [15, 20] and the references therein). This is often called the *Merton problem*, because it was first formulated and solved in a continuous-time setting by Merton [18, 19]. More specifically, he found that – for logarithmic or power utility and one risky asset following geometric Brownian motion – it is optimal to keep the fraction of wealth invested into stocks equal to a constant θ, which is known explicitly in terms of the model parameters.

Magill and Constantinides [17] extended Merton's setting to incorporate proportional transaction costs. In particular, they showed that – again for logarithmic or power utility – it is optimal to engage in the minimal amount of trading necessary to keep the fraction of wealth in stocks inside some *no-trade region* $[\underline{\theta}, \overline{\theta}]$ around θ.

Their somewhat heuristic derivation was made rigorous in the seminal paper of Davis and Norman [4], who also showed how to compute $\underline{\theta}, \overline{\theta}$ by solving a free boundary problem.

Date: August 15, 2011.

2000 *Mathematics Subject Classification.* 91B28, 91B16, 60H10.

Key words and phrases. Transaction costs, optimal consumption, shadow price, asymptotics.

We thank Paolo Guasoni, Mete Soner, and, in particular, Steve Shreve for valuable discussions and comments. We are also grateful to an anonymous referee for his/her careful reading of the manuscript. The first author was partially supported by the Austrian Federal Financing Agency and the Christian-Doppler-Gesellschaft (CDG). The second author gratefully acknowledges partial support by the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK), Project D1 (Mathematical Methods in Financial Risk Management), of the Swiss National Science Foundation (SNF). The third author was partially supported by the Austrian Science Fund (FWF) under grant P19456, the European Research Council (ERC) under grant FA506041, the Vienna Science and Technology Fund (WWTF) under grant MA09-003, and by the Christian-Doppler-Gesellschaft (CDG)..

Using the theory of viscosity solutions, this was further generalized by Shreve and Soner [23]. Moreover, the same approach allowed Janeček and Shreve [12] to derive rigorous first-order asymptotic expansions of $\underline{\theta}, \overline{\theta}$ and the value function for small transaction costs (also cf. [2, 23, 21, 26] for related asymptotic results).

All of the above papers use methods from the theory of stochastic control. On the other hand, dating back to the pioneering papers of Jouini and Kallal [13] and Cvitanić and Karatzas [3], much of the general theory for markets with transaction costs is formulated in terms of duality theory leading to *consistent price systems* resp. *shadow prices*. These are frictionless markets evolving within the bid-ask spread of the original market with transaction costs, which lead to an equivalent dual optimization problem. For logarithmic utility, Kallsen and Muhle-Karbe [14] solved the Merton problem with transaction costs by simultaneously determining a shadow price and its optimal portfolio, i.e., by determining the dual and the primal optimizer at the same time (also compare [16] for an extension of this approach to a limit order market).

In the present article, we revisit this problem using a different parametrization in the spirit of [6]. We employ arguments that are tailor-made for sufficiently small transaction costs, i.e., which focus on asymptotic expansions and thus seem promising for tackling more complicated models. In particular, we determine rigorous asymptotic expansions both for the boundaries $\underline{\theta}, \overline{\theta}$ of the no-trade region and the corresponding value function, where terms of arbitrary order can be algorithmically computed. This extends the first-order expansions of [12]. However, we emphasize that we only consider the particularly simple case of logarithmic utility, unlike [12], who also deal with the more involved case of power utilities. An extension of the present approach to power utilities is subject of current research.

The remainder of this article is organized as follows. After introducing the Merton problem with transaction costs in Section 2, we present a heuristic derivation of our candidate shadow price process. Subsequently, in Section 4, we prove that this candidate is indeed well-defined and a shadow price process for sufficiently small transaction costs. Moreover, we derive asymptotic expansions for the boundaries of the no-trade region. Afterwards, we also determine an asymptotic expansion for the value function. Finally, in Section 6, we show that the shadow price corresponds to the solution of the dual problem.

2. The Merton Problem with transaction costs

We study the problem of maximizing expected logarithmic utility from consumption over an infinite horizon in the presence of proportional transaction costs, as in [4, 14, 23]. Fix a filtered probability space $(\Omega, \mathscr{F}, (\mathscr{F}_t)_{t\in\mathbb{R}_+}, P)$, whose filtration is generated by a standard Brownian motion $(W_t)_{t\in\mathbb{R}_+}$. We consider a market with two investment opportunities, namely a bond and a stock. The *ask price process* S of the stock is supposed to follow the Black-Scholes model

$$dS_t/S_t = \mu dt + \sigma dW_t, \quad S_0 > 0, \quad \mu, \sigma > 0, \tag{2.1}$$

and the *bid price process* is assumed to be given by $(1-\lambda)S$ for some $\lambda \in (0,1)$[1]. This means that one has to pay the higher price S_t when purchasing the stock at

[1]This notation, also used in [25], turns out to be convenient in the sequel. It is equivalent to the usual setup with the same constant proportional transaction costs for purchases and sales (compare, e.g., [4, 12, 14, 23]). Indeed, set $\check{S} = \frac{2-\lambda}{2} S$ and $\check{\lambda} = \frac{\lambda}{2-\lambda}$. Then $((1-\lambda)S, S)$ coincides

time t, but only receives the lower price $(1-\lambda)S_t$ when selling it. The price of the bond is supposed to be constant and equal to $S^0 = 1$. As in [4, 14], we make the following standing assumption, which ensures that the holdings in bond and stock remain positive at all times.

Standing Assumption 2.1.

$$0 < \theta := \frac{\mu}{\sigma^2} < 1.$$

The case $\theta > 1$ can be dealt with using minor modifications of the present approach (cf. [6]). The degenerate case $\theta = 1$ has to be treated separately, see [12, 23].

Since transactions of infinite variation lead to immediate bankruptcy (compare [1]), we confine ourselves to the following set of trading strategies.

Definition 2.2. *A* trading strategy *is an $\mathbb{R}^2$-valued predictable process (φ^0, φ) of finite variation. Here, $(\varphi^0_{0-}, \varphi_{0-}) = (\eta_B, \eta_S) \in \mathbb{R}^2_+ \setminus \{0,0\}$ represents the* initial endowment *in bonds and stocks, whereas φ^0_t and φ_t denote the units held in bond and in stock at time t, respectively. A (discounted)* consumption rate *is an $\mathbb{R}_+$-valued, adapted stochastic process κ satisfying $\int_0^t \kappa_s ds < \infty$ a.s. for all $t \geq 0$. A pair $((\varphi^0, \varphi), \kappa)$ of a trading strategy (φ^0, φ) and a consumption rate κ is called* portfolio/consumption process.

To capture the notion of a self-financing strategy, we use the intuition that no funds are added or withdrawn. As in [14], this leads to the following notion.

Definition 2.3. *A portfolio/consumption process $((\varphi^0, \varphi), \kappa)$ is called* self-financing, *if*

$$\varphi^0 = \varphi^0_{0-} + \int_0^\cdot (1-\lambda) S_t d\varphi^\downarrow_t - \int_0^\cdot S_t d\varphi^\uparrow_t - \int_0^\cdot \kappa_t dt, \tag{2.2}$$

where $\varphi = \varphi^\uparrow - \varphi^\downarrow$ for increasing predictable processes $\varphi^\uparrow, \varphi^\downarrow$ which do not grow at the same time.

Note that since S is continuous and φ is of finite variation, integration by parts yields that this definition coincides with the usual notion of self-financing strategies in the absence of transaction costs if we let $\lambda = 0$.

The subsequent definition requires the investor to be solvent at all times. For frictionless markets, i.e., if $\lambda = 0$, this coincides with the usual notion of non-negativity of wealth.

Definition 2.4. *A self-financing portfolio/consumption process $((\varphi^0, \varphi), \kappa)$ is called* admissible, *if the* liquidation wealth process

$$V_t(\varphi^0, \varphi) := \varphi^0_t + \varphi^+_t (1-\lambda) S_t - \varphi^-_t S_t, \quad t \geq 0,$$

of (φ^0, φ) is a.s. nonnegative.

We consider the following classical investment/consumption problem.

---

with $((1-\check{\lambda})\check{S}, (1+\check{\lambda})\check{S})$. Conversely, any bid-ask process $((1-\check{\lambda})\check{S}, (1+\check{\lambda})\check{S})$ with $\check{\lambda} \in (0,1)$ equals $((1-\lambda)S, S)$ for $S = (1+\check{\lambda})\check{S}$ and $\lambda = \frac{2\check{\lambda}}{1+\check{\lambda}}$.

Definition 2.5 (Primal problem)**.** *Given an initial endowment* (η_B, η_S)*, an admissible portfolio/consumption process* $((\varphi^0, \varphi), \kappa)$ *is called* optimal *if it attains the maximum in*

$$u_{(1-\lambda)S,S}(\eta_B, \eta_S) := \sup_{((\varphi^0,\varphi),\kappa)} E\left(\int_0^\infty e^{-\delta t}\log(\kappa_t)dt\right) \tag{2.3}$$

over all admissible portfolio/consumption processes $((\varphi^0, \varphi), \kappa)$*. Here* $\delta > 0$ *denotes a fixed given* impatience rate*.*

Note that since $\delta > 0$, the value function of the Merton problem *without* transaction costs is finite by [4, Theorem 2.1]. This upper bound for the value function in the present setup with transaction costs then implies that the latter is finite, since it is also bounded from below by the finite utilities that can be obtained from liquidating at time zero and consuming at a sufficiently small constant relative rate.

The dual problem corresponding to (2.3) can be defined as follows (compare [3]).

Definition 2.6 (Dual problem)**.** *Denote by* $\tilde{U}(t, y) = -e^{-\delta t}(\log(e^{\delta t}y) + 1)$ *the* conjugate function *of the utility function* $U(t, x) = e^{-\delta t}\log(x)$*, i.e., the Legendre transform of* $-U(t, -x)$*. Then a pair* $(\tilde{S}, Z)$ *is called* dual optimizer*, if it attains the minimum in*

$$v_{(1-\lambda)S,S}(\eta_B, \eta_S) := \inf_{(\tilde{S},Z)} E\left(\int_0^\infty \tilde{U}\left(t, \frac{Z_t}{\delta(\eta_B + \eta_S \bar{S}_0)}\right) dt\right) + \frac{1}{\delta}. \tag{2.4}$$

Here $(\tilde{S}, Z)$ *runs through the set of all continuous semimartingales* $\tilde{S}$ *with values in the bid-ask spread* $[(1-\lambda)S, S]$ *and the density processes* Z *of corresponding martingale measures, i.e., the positive martingales with initial value* 1 *for which* $\tilde{S}Z$ *is a martingale.*

Remark 2.7. The pairs $(\tilde{S}, Z)$ correspond to the consistent price systems of [9, 22]. In the frictionless case $\lambda = 0$ there is only one of these, namely the stock price process S itself and the corresponding density process Z determined by Girsanov's theorem. In this case, the dual problem reduces to the usual notion for frictionless markets as in, e.g., [15, Theorems 3.9.11 and 3.9.12]. The only difference is the constant $1/\delta$, which is added to let the optimal primal and dual values coincide (compare Theorem 6.1 below).

3. Heuristic derivation of the shadow price

The key to solving both the primal and dual problems above is the following concept:

Definition 3.1. *A* shadow price process $\tilde{S}$ *is a continuous semimartingale with values in the bid-ask spread* $[(1-\lambda)S, S]$*, such that the maximal expected utilities* $u_{\tilde{S},\tilde{S}}(\eta_B, \eta_S)$ *in the frictionless market with price process* $\tilde{S}$ *and* $u_{(1-\lambda)S,S}(\eta_B, \eta_S)$ *in the original market with bid-ask process* $((1-\lambda)S, S)$ *coincide.*

From an economic point of view, it is obvious that the maximal expected utility for *any* frictionless price process $\bar{S}$ with values in the bid-ask spread $[(1-\lambda)S, S]$ is at least as high as for the original problem with transaction costs, since one can always buy at least as cheaply and sell at least as expensively. To achieve equality of the maximal expected utilities, the optimal number φ_t of stocks in the frictionless market with the shadow price process $\tilde{S}$ should increase (resp. decrease) only when

$\tilde{S}_t = S_t$ (resp. $\tilde{S}_t = (1-\lambda)S_t$). Otherwise one could achieve higher utility than in the original market with transaction costs. This portfolio will then also be feasible in the original market with transaction costs and hence be optimal for this market as well.

It is well-known, dating back to [17], that there is no trading in the original market with transaction costs as long as the fraction

$$\pi_t := \frac{\varphi_t S_t}{\varphi_t^0 + \varphi_t S_t} = \frac{1}{1 + \varphi_t^0/(\varphi_t S_t)}$$

of total wealth in terms of the ask price S invested in stocks lies inside some interval $[\underline{\theta}, \overline{\theta}]$ around the Merton proportion $\theta = \mu/\sigma^2$. Let us reparametrize this conveniently for the computations below (compare [6]): Set

$$c := \frac{1-\underline{\theta}}{\underline{\theta}}, \quad \overline{s} := \frac{(1-\underline{\theta})\overline{\theta}}{(1-\overline{\theta})\underline{\theta}} \tag{3.1}$$

and define

$$m_t := \varphi_t^0/(c\varphi_t).$$

Then $\pi \in [\underline{\theta}, \overline{\theta}]$ is equivalent to $m \in [S/\overline{s}, S]$. We now make the *ansatz* that the shadow price $\tilde{S}$ is given by

$$\tilde{S}_t = G(m_t, S_t),$$

for some function $G : \mathbb{R}^2 \to \mathbb{R}$. This is equivalent to assuming that $\tilde{S}_t$ is a function of the current ask price S_t and the current fraction π_t of wealth, in terms of the ask price S, invested in stocks. However, using m_t instead of π_t turns out to be more convenient in the calculations below.

The function G enjoys a scaling property inherited from the logarithmic utility function and its constant *relative* risk aversion:

$$G(am_t, aS_t) = a\tilde{S}_t = aG(m_t, S_t), \quad \forall a > 0.$$

This is because the optimal number φ_t^0 of bonds changes to $a\varphi_t^0$ whereas the optimal number φ_t of stocks and the dimensionless constant c remain the same when we rescale from S_t to aS_t. Therefore, we can pass to the function

$$g(S_t/m_t) := G(1, S_t/m_t),$$

such that

$$\tilde{S}_t = G(m_t, S_t) = m_t g(S_t/m_t).$$

We now heuristically derive a free boundary problem that determines the function g as well as the constants c and $\overline{s}$. With these quantities at hand, we then explain how to construct the process m and in turn the shadow price $\tilde{S}$.

If $\tilde{S}$ is a shadow price, it is well-known that for logarithmic utility the optimal consumption rate at time $t \geq 0$ in the frictionless market with price process $\tilde{S}$ is given by

$$\kappa_t = \delta(\varphi_t^0 + \varphi_t \tilde{S}_t),$$

where $\delta > 0$ denotes the discount factor in the optimization problem (cf., e.g., [20]). Consequently, we have

$$d\varphi_t = 0, \quad d\varphi_t^0 = -\delta(\varphi_t^0 + \varphi_t \tilde{S}_t)dt = -\delta\varphi_t^0(1 + g(S_t/m_t)/c)dt \tag{3.2}$$

in the no-trade region, where we have inserted our ansatz $\tilde{S}_t = m_t g(S_t/m_t)$ and the definition $m_t = \varphi_t^0/(c\varphi_t)$. Now consider the process

$$Y_t := S_t/m_t = S_t c\varphi_t/\varphi_t^0,$$

which represents the fraction of wealth in stocks (in terms of the ask price S) over wealth in bonds, normalized by c. In view of the dynamics (2.1) of S and (3.2) of φ^0 and φ, Itô's formula yields

$$dY_t/Y_t = dS_t/S_t - d\varphi_t^0/\varphi_t^0 = \big(\mu + \delta(1+g(Y_t)/c)\big)dt + \sigma dW_t. \tag{3.3}$$

Again by Itô's formula, the dynamics of the shadow price $\tilde{S} = (S/Y)g(Y)$ are

$$\begin{aligned} d\tilde{S}_t/\tilde{S}_t &= \left(\frac{\mu g'(Y_t)Y_t + \frac{\sigma^2}{2}g''(Y_t)Y_t^2 - \delta(g(Y_t) - g'(Y_t)Y_t)(1+g(Y_t)/c)}{g(Y_t)}\right)dt \\ &\quad + \left(\frac{\sigma g'(Y_t)Y_t}{g(Y_t)}\right)dW_t \\ &=: \tilde{\mu}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t, \end{aligned}$$

during each excursion into the no-trade region. Due to Merton's rule for logarithmic utility [19], the Merton ratio for $\tilde{S}$ must equal the proportion of total wealth invested in stocks *in terms of the shadow price* $\tilde{S}$, i.e., we must have

$$\frac{\tilde{\mu}(Y_t)}{\tilde{\sigma}(Y_t)^2} = \tilde{\pi}_t := \frac{\varphi_t\tilde{S}_t}{\varphi_t^0 + \varphi_t\tilde{S}_t} = \frac{g(Y_t)}{c+g(Y_t)}.$$

After rearranging, this yields the following ODE for g:

$$g''(y) = \frac{2g'(y)^2}{c+g(y)} - \frac{2\theta g'(y)}{y} + \frac{2\delta}{\sigma^2}\left(1+\frac{g(y)}{c}\right)\left(\frac{g(y)}{y^2} - \frac{g'(y)}{y}\right). \tag{3.4}$$

Note that in the limit $\delta \to 0$, this is precisely the ODE from [6]. By (3.1), the process $Y = S/m$ takes values in $[1, \overline{s}]$. Hence it is natural to complement the ODE (3.4) with some boundary conditions at 1 and $\overline{s}$ as in [6]. Since we have a second order ODE, and the parameters c and $\overline{s}$ are also yet to be determined, we need two boundary conditions at each endpoint. The shadow price $\tilde{S} = mg(S/m)$ should equal the bid price $(1-\lambda)S$ at the selling boundary $\overline{\theta}$, where $m = S/\overline{s}$, resp. the ask price S at the buying boundary $\underline{\theta}$, where $m = S$. Hence we impose

$$g(1) = 1 \quad \text{and} \quad g(\overline{s}) = (1-\lambda)\overline{s}. \tag{3.5}$$

Moreover, we add the smooth pasting conditions

$$g'(1) = 1 \quad \text{and} \quad g'(\overline{s}) = 1-\lambda. \tag{3.6}$$

They can be derived by looking at the ratio $\tilde{S}/S$, which should stay in $[1-\lambda, 1]$, and setting its diffusion coefficient equal to zero at $1-\lambda$ and 1 (compare [6]). These four boundary conditions should now determine the function g as the solution to our second-order ODE as well as the two free variables c and $\overline{s}$, which in turn identify the boundaries $\underline{\theta}$ and $\overline{\theta}$ of the no-trade region.

With $g, c, \overline{s}$ at hand, it now remains to extend this construction from one excursion into the no-trade region to the entire positive real time line. To this end notice that, even when trading in the stock does happen, the process $Y = c\varphi_t S_t/\varphi_t^0$ must always remain within $[1, \overline{s}]$, in order to keep the corresponding fraction of wealth in stocks in $[\underline{\theta}, \overline{\theta}]$. However, the diffusion coefficient in the SDE (3.3) is bounded away from zero when Y_t approaches either 1 or $\overline{s}$. Therefore we need to complement the SDE (3.3) for Y with *instantaneous reflection* at both endpoints of $[1, \overline{s}]$. Then Y will by definition remain in $[1, \overline{s}]$ and follow the claimed dynamics during each excursion into the no-trade region. With the process Y at hand, we can then define the candidate shadow price $\tilde{S} = (S/Y)g(Y)$.

4. Construction of the shadow price

We begin with an existence result for the free boundary value problem from the preceding section. As in [14, Proposition 4.2], it would be possible to obtain a similar result for any relative width λ of the bid-ask spread. However, we restrict ourselves to the case of sufficiently small λ here. This is all we need to derive asymptotic expansions in terms of λ below and seems much better suited to the analysis of more complicated models, since it requires only a much coarser inspection of the involved equations.

Lemma 4.1. *If $\lambda > 0$ is sufficiently small, then there are constants $c > (1-\theta)/\theta$ and $\bar{s} > 1$ as well as a function $g : [1, \bar{s}] \to [1, (1-\lambda)\bar{s}]$ satisfying the ODE*

$$g''(y) = \frac{2g'(y)^2}{c+g(y)} - \frac{2\theta g'(y)}{y} + \frac{2\delta}{\sigma^2}\left(1 + \frac{g(y)}{c}\right)\left(\frac{g(y)}{y^2} - \frac{g'(y)}{y}\right) \tag{4.1}$$

with boundary conditions

$$g(1) = 1, \quad g(\bar{s}) = (1-\lambda)\bar{s}, \quad g'(1) = 1, \quad g'(\bar{s}) = (1-\lambda). \tag{4.2}$$

Proof. We first ignore the free boundary $\bar{s}$ and consider only the ODE (4.1) with initial conditions $g(1) = g'(1) = 1$, with c as free parameter. This problem can be written as

$$\begin{aligned} \partial_s g(s,c) &= F_1(s,g,h,c), \\ \partial_s h(s,c) &= F_2(s,g,h,c), \\ g(1,c) &= h(1,c) = 1, \end{aligned}$$

where $F_1(s,u,v,c) := v$, and

$$F_2(s,u,v,c) := \frac{2v^2}{c+u} - \frac{2\theta v}{s} + \frac{2\delta}{\sigma^2}\left(1 + \frac{u}{c}\right)\left(\frac{u}{s^2} - \frac{v}{s}\right).$$

Since $\bar{c} := (1-\theta)/\theta \notin \{-1, 0\}$, the function F_2 (and of course also F_1) is analytic in a complex neighbourhood of $(s,u,v,c) = (1,1,1,\bar{c})$. By a standard existence and uniqueness result for analytic ODEs [11, Theorem 1.1], we obtain a neighbourhood of $(s,c) = (1,\bar{c})$ with a unique analytic solution $g(s,c)$. The coefficients a_{ij} of its Taylor expansion

$$g(s,c) = 1 + (s-1) + \sum_{i \geq 2} \sum_{j \geq 0} a_{ij} (s-1)^i (c-\bar{c})^j$$

can be calculated recursively. Note that $a_{0j} = a_{1j} = \delta_{0j}$ for $j \geq 0$, where δ is the Kronecker delta, which follows from the initial conditions $g(1,c) = g'(1,c) = 1$.

From this analytic function of two variables we will now construct the solution of the free boundary value problem (4.1)–(4.2), by supplying appropriate quantities c and $\bar{s}$ (depending on λ). These should satisfy

$$\lambda\bar{s} = \bar{s} - g(\bar{s}, c) \qquad \text{and} \qquad g(\bar{s}, c) - \bar{s}\partial_s|_{s=\bar{s}} g(s,c) = 0. \tag{4.3}$$

We divide the latter equation by $s-1$, reflecting the fact that we are not interested in the solution $s = 1$ (which exists because of the initial conditions $g(1) = g'(1) = 1$). A Taylor expansion yields

$$\frac{g(s,c) - s\partial_s g(s,c)}{s-1} = \sum_{i \geq 0} \sum_{j \geq 0} b_{ij} (s-1)^i (c-\bar{c})^j \tag{4.4}$$

for some real, explicitly computable coefficients b_{ij}. It turns out that $b_{0,0} = 0$, while $b_{1,0} = \theta(\theta - 1)$ is non-zero, so that we can apply the implicit function theorem for multivariate analytic functions [10, Theorem I.B.4] to obtain an analytic function

$$H(c) = 1 + \sum_{i \geq 1} d_i (c - \bar{c})^i = 1 + \tfrac{2\theta}{1-\theta}(c - \bar{c}) + O((c - \bar{c})^2) \tag{4.5}$$

such that (4.4) vanishes if $H(c)$ is substituted for s. Here and in what follows, the coefficients of the power series symbolized by the $O(\cdot)$-s can be algorithmically computed. Now we insert the function $H(c)$ into the right hand side of the first equation in (4.3):

$$H(c) - g(H(c), c) = a_0^{-1}(c - \bar{c})^3 + O((c - \bar{c})^4),$$

where $a_0 = \frac{3}{4}(1 - \theta)^2/\theta^4$. If we insert $H(c)$ also on the left hand side, the first equation in (4.3) becomes

$$\lambda(1 + d_1(c - \bar{c}) + O((c - \bar{c})^2)) = a_0^{-1}(c - \bar{c})^3(1 + O(c - \bar{c})).$$

Dividing by $a_0^{-1}(1 + O(c - \bar{c}))$, we obtain

$$\lambda(a_0 + O(c - \bar{c})) = (c - \bar{c})^3.$$

Let us denote the coefficients of the series on the left hand side by a_k:

$$\lambda(a_0 + a_1(c - \bar{c}) + \dots) = (c - \bar{c})^3.$$

Now take the third root:

$$\lambda^{1/3} a_0^{1/3}(1 + \tfrac{a_1}{3a_0}(c - \bar{c}) + \dots) = c - \bar{c}.$$

(Note that the coefficients of powers of power series can be algorithmically calculated [8].) Since $a_0 = \frac{3}{4}(1 - \theta)^2/\theta^4$ is non-zero, we can apply the implicit function theorem, so that the latter equation has a unique solution for small λ. It is an analytic function of $\lambda^{1/3}$:

$$c - \bar{c} = a_0^{1/3}\lambda^{1/3} + O(\lambda^{2/3}).$$

The coefficients in this power series expansion can be calculated with the Lagrange inversion formula:

$$[\lambda^{k/3}](c - \bar{c}) = \tfrac{1}{k}[z^{k-1}] a_0^{k/3}(1 + \tfrac{a_1}{3a_0} z + \dots)^k, \qquad k \geq 1, \tag{4.6}$$

where the operator $[z^k]$ extracts the k-th coefficient of a power series. (This is analogous to the proof of Proposition 6.1 in [6].) With this value of c, the function $g(s, c)$, together with $\bar{s} = H(c)$, is the desired solution of the free boundary value problem. To see that it maps $[1, \bar{s}]$ to $[1, (1 - \lambda)\bar{s}]$, note that otherwise the equation $g(s) - sg'(s) = 0$ would have a solution in the open interval $]1, \bar{s}[$. This is impossible, since $H(c)$ is the *unique* zero of (4.4).

From (4.6), we can calculate as many terms in the expansion of c as we want. The first of them are

$$c = \bar{c} + \frac{1 - \theta}{2\theta}\left(\frac{6}{\theta(1 - \theta)}\right)^{1/3}\lambda^{1/3} + \frac{2\delta/\sigma^2 + 3(1 - \theta)^2}{12\theta}\left(\frac{6}{\theta(1 - \theta)}\right)^{2/3}\lambda^{2/3} + O(\lambda). \tag{4.7}$$

Plugging this expansion into (4.5) gives an expansion of $\bar{s}$, again to arbitrary order. The first terms are

$$\bar{s} = 1 + \left(\frac{6}{\theta(1-\theta)}\right)^{1/3} \lambda^{1/3} + \frac{1}{2}\left(\frac{6}{\theta(1-\theta)}\right)^{2/3} \lambda^{2/3} + O(\lambda). \tag{4.8}$$

Finally, note that (4.7) shows that $c > \bar{c} = (1-\theta)/\theta$ for sufficiently small λ. □

Henceforth, we consider sufficiently small transaction costs $\lambda > 0$ for Lemma 4.1 to be valid and write g resp. $c, \bar{s}$ for the corresponding function resp. constants defined therein.

The existence of the process Y as the solution to the SDE (3.3) with reflection at 1 and $\bar{s}$ is a consequence of a classical result of Skorokhod [24].

Lemma 4.2. *Let $y \in [1, \bar{s}]$. Then there exists a strong solution to the* Skorokhod SDE

$$dY_t = Y_t\big(\mu + \delta(1 + g(Y_t)/c)\big)dt + Y_t \sigma dW_t, \quad Y_0 = y, \tag{4.9}$$

with instantaneous reflection at 1 and $\bar{s}$, that is, a continuous, adapted, $[1, \bar{s}]$-valued process Y and nondecreasing adapted processes L and U increasing only on the sets $\{Y_t = 1\}$ and $\{Y_t = \bar{s}\}$, respectively, such that

$$Y_t = Y_0 + \int_0^t Y_u(\mu + \delta(1 + g(Y_u)/c)du + \int_0^t Y_u \sigma dW_u + \int_0^t Y_u dL_u - \int_0^t Y_u dU_u$$

holds for all $t \in \mathbb{R}_+$.

Remark 4.3. In the limit $\delta \to 0$, the process Y is just a geometric Brownian motion instantaneously reflected at 1 and $\bar{s}$. Hence it coincides with the process S/m from [6]. Let us also point out that Y is related to the process β from [14, Lemma 4.3] by a nonlinear transformation. However, the present parametrization evidently leads to simpler dynamics. Moreover, the additional process C from [14, Lemma 4.4] is no longer necessary.

Proof of Lemma 4.2. Since g is bounded on $[1, \bar{s}]$ and $c > (1-\theta)/\theta > 0$ by Lemma 4.1, the coefficients of (4.9) are globally Lipschitz on $[1, \bar{s}]$. Hence the claim follows from [24]. □

It was shown in [6, Proposition 2.1] that, for $\delta = 0$, the process $\tilde{S} = mg(S/m) = (S/Y)g(Y)$ is an Itô process, i.e., does not involve local time. We now deduce from Lemma 4.2 that this still holds for arbitrary $\delta > 0$. The present proof only uses Itô's formula and the smooth pasting conditions (4.2) for g. In particular, it allows to avoid the somewhat ad hoc localization procedures used in the proof of [6, Proposition 2.1].

Proposition 4.4. *For $y \in [1, \bar{s}]$ and Y as in Lemma 4.2, $\tilde{S} := (S/Y)g(Y)$ is a positive Itô process of the exponential form*

$$\tilde{S} = \frac{S_0 g(y)}{y} \mathscr{E}(\tilde{X}),$$

with

$$d\tilde{X}_t := d\tilde{S}_t/\tilde{S}_t = \left(\frac{\sigma^2 g'(Y_t)^2 Y_t^2}{(c + g(Y_t))g(Y_t)}\right) dt + \left(\frac{g'(Y_t) Y_t \sigma}{g(Y_t)}\right) dW_t, \quad \tilde{X}_0 = 0,$$

and $\tilde{S}$ takes values in the bid-ask spread $[(1-\lambda)S, S]$.

Proof. Using Itô's formula, we obtain

$$dg(Y_t) = g'(Y_t)Y_t(\mu + \delta(1 + g(Y_t)/c)dt + \sigma dW_t + dL_t - dU_t) + \frac{\sigma^2}{2}g''(Y_t)Y_t^2 dt$$

and

$$\frac{d(S_t/Y_t)}{S_t/Y_t} = -\delta\left(1 + \frac{g(Y_t)}{c}\right)dt - dL_t + dU_t.$$

Hence integration by parts yields

$$\begin{aligned} d\tilde{S}_t/\tilde{S}_t = &\left(\frac{\mu g'(Y_t)Y_t + \frac{\sigma^2}{2}g''(Y_t)Y_t^2 - \delta(g(Y_t) - g'(Y_t)Y_t)(1 + g(Y_t)/c)}{g(Y_t)}\right)dt \\ &+ \left(\frac{\sigma g'(Y_t)Y_t}{g(Y_t)}\right)dW_t + \frac{g'(Y_t)Y_t - g(Y_t)}{g(Y_t)}(dL_t - dU_t). \end{aligned}$$

Since L and U only increase on $\{Y_t = 1\}$ resp. $\{Y_t = \overline{s}\}$, where the integrand $g'(Y_t)Y_t - g(Y_t)$ is zero by the smooth pasting conditions for g, the last term vanishes. The Itô process decomposition of $\tilde{X}$ now follows by inserting the ODE (4.1) for g''.

For the second part of the assertion, notice that $g''(1) < 0$, because $c > (1-\theta)/\theta$. Hence it follows from a comparison argument that $g'(y)y - g(y) \leq 0$ on $[1, \overline{s}]$. As this implies that the derivative $(g'(y)y - g(y))/y^2$ of $g(y)/y$ is non-positive, $g(1)/1 = 1$ and $g(\overline{s})/\overline{s} = 1 - \lambda$ yield that $\tilde{S} = Sg(Y)/Y$ is indeed $[(1-\lambda)S, S]$-valued. □

Using standard results for frictionless markets, we can now determine the optimal portfolio/consumption pair for $\tilde{S} = (S/Y)g(Y)$.

Lemma 4.5. *Set*

$$y = \begin{cases} 1, & \text{if } \eta_B > S_0 c\eta_S, \\ \overline{s}, & \text{if } \eta_B < S_0 c\eta_S/\overline{s}, \\ S_0 c\eta_S/\eta_B & \text{otherwise.} \end{cases} \tag{4.10}$$

For Y and $\tilde{S}$ as in Lemma 4.2 and Proposition 4.4, respectively, define

$$\tilde{\pi}_t = \frac{1}{1 + c/g(Y_t)},$$

as well as

$$\begin{aligned} \tilde{V}_t &= (\eta_B + \eta_S\tilde{S}_0)\mathscr{E}\left(\int_0^\cdot \frac{\tilde{\pi}_t}{\tilde{S}_t}d\tilde{S}_t - \int_0^\cdot \delta dt\right), \\ \kappa_t &= \delta\tilde{V}_t, \\ \varphi_t &= \tilde{\pi}_t\tilde{V}_t/\tilde{S}_t, \qquad \varphi_t^0 = (1 - \tilde{\pi}_t)\tilde{V}_t. \end{aligned}$$

Then

$$\varphi_t = \varphi_0 + \int_0^t \frac{\varphi_u c g'(Y_u)Y_u}{g(Y_u)(c + g(Y_u))}dL_u - \int_0^t \frac{\varphi_u c g'(Y_u)Y_u}{g(Y_u)(c + g(Y_u))}dU_u, \tag{4.11}$$

and $((\varphi^0, \varphi), \kappa)$ is an optimal portfolio/consumption pair for the frictionless market with price process $\tilde{S}$. The corresponding wealth process and the optimal fraction of wealth invested in stocks in terms of $\tilde{S}$ are given by $\tilde{V}$ and $\tilde{\pi}$, respectively.

Note that the first resp. second cases in (4.10) occur if the initial fraction of wealth in stock $\eta_S S_0/(\eta_B+\eta_S S_0)$ lies below $\underline{\theta} = (1+c)^{-1}$ resp. above $\overline{\theta} = (1+c/\overline{s})^{-1}$. In this case, there is a jump from the initial position $(\varphi^0_{0-}, \varphi_{0-}) = (\eta_B, \eta_S)$, which moves the initial fraction of wealth to the nearest boundary of the interval $[\underline{\theta}, \overline{\theta}]$. Since this initial bulk trade involves the puchase resp. sale of stocks, the initial value of $\tilde{S}$ is chosen so as to match the initial ask resp. bid price in this case.

Proof. Since $d\tilde{S}/\tilde{S} = d\tilde{X}_t := \tilde{\mu}(Y_t)dt + \tilde{\sigma}(Y_t)dW_t$ is an Itô process with bounded coefficients and $\tilde{\sigma}$ is bounded away from zero, its optimal portfolio/consumption pair is characterized by standard results for frictionless markets, cf., e.g., [7, Theorem 3.1]. In particular, the optimal fraction of wealth (in terms of $\tilde{S}$) invested in stocks is given by Merton's rule, i.e., equals

$$\frac{\tilde{\mu}(Y_t)}{\tilde{\sigma}(Y_t)^2} = \frac{g(Y_t)}{c+g(Y_t)} = \tilde{\pi}_t$$

as claimed. Moreover, the optimal wealth process, the optimal consumption rate and the optimal numbers of bonds resp. stocks are indeed given by $\tilde{V}$, κ, φ^0, and φ. Now Itô's formula and the ODE (4.1) for g imply

$$d\tilde{\pi}_t = \left(\frac{\delta g(Y_t)}{c+g(Y_t)}\right)dt + \left(\frac{c\sigma g'(Y_t)Y_t}{(c+g(Y_t))^2}\right)dW_t + \left(\frac{cg'(Y_t)Y_t}{(c+g(Y_t))^2}\right)(dL_t - dU_t).$$

Moreover, by Yor's formula, we have

$$\begin{aligned}
&\frac{\mathscr{E}(\int_0^\cdot \frac{g(Y_t)}{c+g(Y_t)}d\tilde{X}_t - \int_0^\cdot \delta dt)}{\mathscr{E}(\tilde{X})} \\
&\quad= \mathscr{E}\left(\int_0^\cdot \frac{c}{c+g(Y_t)}d(\langle\tilde{X},\tilde{X}\rangle_t - \tilde{X}_t) - \int_0^\cdot \delta dt\right) \\
&\quad= \mathscr{E}\left(\int_0^\cdot \frac{c^2\sigma^2 g'(Y_t)^2 Y_t^2}{g(Y_t)^2(c+g(Y_t))^2}dt - \int_0^\cdot \delta dt - \int_0^\cdot \frac{c\sigma g'(Y_t)Y_t}{g(Y_t)(c+g(Y_t))}dW_t\right).
\end{aligned}$$

Taking into account that the process $L-U$ is continuous and of finite variation, integration by parts finally yields the claimed representation for φ. □

The representation (4.11) and the definition of $\tilde{S}_0$ via (4.10) imply that stocks are only purchased (resp. sold) when $\tilde{S} = S$ (resp. $\tilde{S} = (1-\lambda)S$). Hence the optimal portfolio/consumption process $((\varphi^0, \varphi), \kappa)$ for $\tilde{S}$ is also admissible for the bid-ask process $((1-\lambda)S, S)$. Since shares can be bought at least as cheaply and sold as least as expensively when trading $\tilde{S}$ instead of $(1-\lambda)S, S)$, it follows that the maximal expected utilities coincide, i.e., $\tilde{S}$ is a shadow price. Made precise, this is the content of the following analogue of [14, Theorem 4.6].

Theorem 4.6. *The portfolio/consumption process $((\varphi^0, \varphi), \kappa)$ from Lemma 4.5 is also optimal for the bid-ask process $((1-\lambda)S, S)$. In particular, $\tilde{S}$ is a shadow price in this market.*

Proof. This follows verbatim as in the proof of [14, Theorem 4.6]. □

In view of Lemma 4.5 and the definition of g, the optimal proportion $\tilde{\pi}$ of stocks in terms of the shadow price $\tilde{S}$ takes values in the interval $[(1+c)^{-1}, (1+c/((1-\lambda)\overline{s})^{-1}]$. Translating this into terms of the ask price S, we find that the optimal fraction π in terms of S is kept within $[(1+c)^{-1}, (1+c/\overline{s})^{-1}]$, i.e., the lower resp. upper

boundaries of the no-trade region are given by $\underline{\theta} = (1+c)^{-1}$ resp. $\overline{\theta} = (1+c/\overline{s})^{-1}$. The following series expansions are immediate consequences of (4.7) and (4.8) (upon taking one additional term in both expansions).

Corollary 4.7. *The lower and upper boundaries of the no-trade region in terms of the ask price S have the expansions*

$$\begin{aligned}\underline{\theta} = \frac{1}{1+c} &= \theta - \left(\frac{3}{4}\theta^2(1-\theta)^2\right)^{1/3}\lambda^{1/3} - \frac{\delta\theta}{6\sigma^2}\left(\frac{6}{\theta(1-\theta)}\right)^{2/3}\lambda^{2/3} \\ &\quad + \frac{\theta(9(1-\theta)^2(1-2\theta+2\theta^2) - 4\delta^2/\sigma^4 + 6\delta/\sigma^2(3-8\theta+5\theta^2))}{360(1-\theta)}\frac{6}{\theta(1-\theta)}\lambda \\ &\quad + O(\lambda^{4/3})\end{aligned}$$

and

$$\begin{aligned}\overline{\theta} = \frac{1}{1+c/\overline{s}} &= \theta + \left(\frac{3}{4}\theta^2(1-\theta)^2\right)^{1/3}\lambda^{1/3} - \frac{\delta\theta}{6\sigma^2}\left(\frac{6}{\theta(1-\theta)}\right)^{2/3}\lambda^{2/3} \\ &\quad - \frac{\theta(3(1-\theta)^2(3-26\theta+26\theta^2) - 4\delta^2/\sigma^4 + 6\delta/\sigma^2(3-8\theta+5\theta^2))}{360(1-\theta)}\frac{6}{\theta(1-\theta)}\lambda \\ &\quad + O(\lambda^{4/3}),\end{aligned}$$

respectively. Consequently, the size of the no-trade region in terms of S satisfies

$$\begin{aligned}\overline{\theta}-\underline{\theta} &= \frac{1}{1+c/\overline{s}} - \frac{1}{1+c} \\ &= (6\theta^2(1-\theta)^2)^{1/3}\lambda^{1/3} \\ &\quad - \frac{\theta(3(1-\theta)^2(3-16\theta+16\theta^2) - 4\delta^2/\sigma^4 + 6\delta/\sigma^2(3-8\theta+5\theta^2))}{180(1-\theta)}\frac{6}{\theta(1-\theta)}\lambda \\ &\quad + O(\lambda^{4/3}).\end{aligned}$$

Further terms of these expansions can be algorithmically computed, if desired.

The first correction terms are of order $\lambda^{1/3}$, which has been conjectured in [23, Remark B.3] and proved in [12]. Whereas the first-order corrections are symmetric around the Merton proportion θ, the second correction terms of order $\lambda^{2/3}$ are both negative. This means that it is optimal to start transacting for smaller fractions of stock than indicated by the first-order expansions. As pointed out by [12], an intuitive explanation is that consumption from the bank account continuously increases the proportion of stock. Accordingly, the $\lambda^{2/3}$-terms vanish in the limit for $\delta \to 0$. Interestingly, however, the impatience rate $\delta > 0$ only shows up starting from the third term of the expansion for the *width* of the no-trade region even in the presence of consumption.

Remark 4.8. We now compare the expansions from Corollary 4.7 to the results of Janeček and Shreve [12] mentioned above. We rename their parameter λ as $\check{\lambda}$; it is related to our λ by $\check{\lambda} = \lambda/(2-\lambda)$ (see the footnote in Section 2). The bounds of the no-trade region become

(4.12)

$$\frac{1}{1+2c/(2-\check{\lambda})} = \theta - \left(\frac{3}{2}\theta^2(1-\theta)^2\right)^{1/3}\check{\lambda}^{1/3} - \frac{2\theta\delta}{\sigma^2}(12\theta^2(1-\theta)^2)^{-1/3}\check{\lambda}^{2/3} + O(\check{\lambda})$$

and

(4.13)
$$\frac{1}{1+2c/((2-\check{\lambda})\bar{s})} = \theta + \left(\frac{3}{2}\theta^2(1-\theta)^2\right)^{1/3}\check{\lambda}^{1/3} - \frac{2\theta\delta}{\sigma^2}(12\theta^2(1-\theta)^2)^{-1/3}\check{\lambda}^{2/3} + O(\check{\lambda}).$$

The $\check{\lambda}^{1/3}$-terms agree with those established in [12], while our $\check{\lambda}^{2/3}$-terms are not those that were conjectured in [12], based on heuristic considerations.

We are indebted to Steve Shreve, who kindly checked these calculations again. While the arguments worked out in [12] turned out to be perfectly correct, he found that an assumption on which these calculations were based did not hold true. It was stated in equation (4.27) of [12] that "there is considerable evidence" that the coefficient of the λ term in the expansion of the value function has a certain nonzero value, while subsequent calculations by Steve Shreve indicate that this coefficient is zero. When this term is set to zero, then heuristic calculations based on the approach of [12] lead to (4.12) and (4.13) as well.

5. Characterization and asymptotic expansion of the value function

In the frictionless market with price process $\tilde{S}$, standard control methods allow to characterize the value function for the primal problem from Definition 2.4 as the unique classical solution to the corresponding HJB equation. This yields the maximal utility in the frictionless shadow market, which coincides with the maximal utility in the original market with bid-ask process $((1-\lambda)S, S)$ by Theorem 4.6.

Theorem 5.1. *Let y, Y, and $\tilde{S}$ be defined as in Lemma 4.5. Then, if $\lambda > 0$ is sufficiently small, we have*

$$u_{(1-\lambda)S,S}(\eta_B, \eta_S) = \log(\eta_B + \eta_S\tilde{S}_0)/\delta + w(y), \tag{5.1}$$

where w is the unique (classical) solution to the ODE

$$\frac{\sigma^2}{2}s^2w''(s) + \left(\mu + \delta\left(1 + \frac{g(s)}{c}\right)\right)sw'(s) - \delta w(s) + \frac{\tilde{\mu}^2(s)}{2\delta\tilde{\sigma}^2(s)} + \log(\delta) - 1 = 0, \tag{5.2}$$

on $[1, \bar{s}]$ with Neumann boundary conditions

$$w'(1) = w'(\bar{s}) = 0.$$

Proof. We first show that the boundary value problem (5.2) has a unique solution, which is an analytic function. This is similar to the first part of the proof of Lemma 4.1. We first solve the ODE (5.2) with initial conditions $w(1) = w_1 + \xi$ and $w'(1) = 0$, with arbitrary ξ, and where

$$w_1 = \frac{\mu^2 - 2\delta\sigma^2}{2\delta^2\sigma^2} + \frac{\log\delta}{\delta}$$

is the value to which $w(1)$ converges for $\lambda \to 0$. This yields an analytic function

$$w(s, c, \xi) = w_1 + \xi + \sum_{i\geq 2}\sum_{j,k\geq 0} a_{ijk}(s-1)^i\xi^j(c-\bar{c})^k,$$

for certain real coefficients a_{ijk}. (Recall that c tends to $\bar{c} = (1-\theta)/\theta$ as $\lambda \to 0$.) To satisfy the boundary condition $w'(\bar{s}) = 0$, we solve the equation

$$\frac{\partial}{\partial s}w(s, c, \xi)\bigg|_{s=\bar{s}(c)} = 0$$

for ξ. After inserting the expansion (4.5) for $\bar{s}$, we obtain a solution $\xi(c)$, analytic at $c = \bar{c}$. Then $w(s) = w(s, \xi(c(\lambda)), c(\lambda))$ solves the boundary value problem.

Now fix $\lambda > 0$ sufficiently small, such that Lemma 4.1, Corollary 4.7, and the above existence result for w can be applied. Then, in particular, the mapping $y \mapsto w(y)$ can be extended to a C^2-mapping on the whole real line, e.g., by attaching suitable parabolas at $y = 1$ and $y = \bar{s}$. Itô's formula in turn shows that

$$\begin{aligned} M_t :=& \int_0^t e^{-\delta u} \log(\kappa_u) du + e^{-\delta t} \left(\frac{\log(\tilde{V}_t)}{\delta} + w(Y_t) \right) \\ =& M_0 + \int_0^t e^{-\delta u} \left(\frac{\tilde{\sigma}(Y_u)\tilde{\pi}_u}{\delta} + \sigma Y_u w'(Y_u) \right) dW_u, \end{aligned}$$

where we have inserted the dynamics of $\tilde{V}$ (cf. Lemma 4.5) and Y (cf. Lemma 4.2) as well as the ODE and boundary conditions for w. In particular, the process $(M_t)_{t \geq 0}$ is a local martingale. By the proof of Proposition 4.4, we have $\tilde{\sigma}(y) = \sigma g'(y)/(g(y)/y) \leq \sigma/(1-\lambda)$ for $y \in [1, \bar{s}]$, such that $0 \leq \tilde{\pi} \leq (1 + c/((1-\lambda)\bar{s}))^{-1} \leq (1 + c/\bar{s})^{-1}$ and $1 \leq Y \leq \bar{s}$ imply

$$\int_0^t E\left[e^{-2\delta u} \left(\frac{\tilde{\sigma}(Y_u)\tilde{\pi}_u}{\delta} + \sigma Y_u w'(Y_u) \right)^2 \right] du \leq \frac{1}{2\delta} \left(\frac{\sigma/(1-\lambda)}{\delta(1 + c/\bar{s})} + \sigma \bar{s} |w'|_{\max} \right)^2,$$

for $|w'|_{\max} := \max_{y \in [1, \bar{s}]} |w'(y)|$. Therefore the process $(M_t)_{t \geq 0}$ is in fact a true martingale and also uniformly integrable, since it is bounded in $L^2(P)$ by the Itô isometry. Hence M_t converges in $L^1(P)$ to a random variable M_∞ with

$$E(M_\infty) = E(M_0) = \log(\tilde{V}_0)/\delta + w(Y_0) = \log(\eta_B + \eta_S \tilde{S}_0)/\delta + w(y). \tag{5.3}$$

Now notice that $e^{-\delta t} w(Y_t) \to 0$, because $w(\cdot)$ is bounded on $[1, \bar{s}]$. Moreover, as $\log(\tilde{V}_t)$ is an Itô process with bounded coefficients, it follows that $e^{-\delta t} \log(\tilde{V}_t)/\delta \to 0$ in $L^2(P)$ and hence also in probability. Therefore $M_t \to \int_0^\infty e^{-\delta t} \log(\kappa_t) dt$ in probability, such that the a.s. uniqueness of limits in probability yields

$$E\left(\int_0^\infty e^{-\delta t} \log(\kappa_t) dt \right) = E(M_\infty).$$

Combined with (5.3), this completes the proof. □

It does not seem possible to come up with an explicit solution to the above ODE even in terms of the function g, unlike in the case without consumption treated in [5, 25]. However, we can again derive an expansion in powers of $\lambda^{1/3}$. Note that this expansion cannot be recovered by letting the relative risk aversion tend to 1 in the corresponding result for power utility [12]. The reason is that the value function for power utility does not converge to the value function for log utility.

Corollary 5.2. *In addition to the assumptions of Theorem 5.1, suppose that our portfolio starts on the Merton line, i.e.,*

$$\pi_0 = \frac{\eta_S S_0}{\eta_B + \eta_S S_0} = \theta.$$

Then the value function satisfies

$$u_{(1-\lambda)S,S}(\eta_B,\eta_S) = \frac{1}{\delta}\log(\delta(\eta_B+\eta_S S_0)) + \frac{1}{\delta^2}\left(\frac{\mu^2}{2\sigma^2}-\delta\right) - \frac{3^{2/3}(\mu(\sigma^2-\mu))^{4/3}}{4\cdot 2^{1/3}\delta^2\sigma^{10/3}}\lambda^{2/3}+O(\lambda). \tag{5.4}$$

This expansion can be continued to any order, with algorithmically computable coefficients.

Proof. We expand the first term in (5.1):

$$\begin{aligned}\frac{1}{\delta}\log(\eta_B+\eta_S\tilde{S}_0) &= \frac{1}{\delta}\log(\eta_B+\eta_S S_0 g(y)/y)\\ &= \frac{1}{\delta}\log(\eta_B+\eta_S S_0) + \frac{1}{\delta}\log\left(1+\theta\left(\frac{g(y)}{y}-1\right)\right)\\ &= \frac{1}{\delta}\log(\eta_B+\eta_S S_0)+O(\lambda).\end{aligned}$$

The last line follows after expanding log into a Taylor series, and inserting $y = c/\bar{c}$ and the series for g and c from the proof of Lemma 4.1. As for the second term in (5.1), we use the series for w found in the proof of Theorem 5.1 and obtain

$$w(y) = \frac{1}{\delta^2}\left(\frac{\mu^2}{2\sigma^2}-\delta\right) + \frac{\log\delta}{\delta} - \frac{3^{2/3}(\mu(\sigma^2-\mu))^{4/3}}{4\cdot 2^{1/3}\delta^2\sigma^{10/3}}\lambda^{2/3}+O(\lambda).$$

Adding both expansions yields (5.4). □

Note that the expansion (5.4) can easily be refined to a bivariate expansion in terms of $\lambda^{1/3}$ and $\pi_0-\theta$, valid for $\lambda>0$ sufficiently small and $\pi_0\in[\underline{\theta},\overline{\theta}]$ inside the no-trade region. Its first terms are identical to (5.4); $\pi_0-\theta$ only appears from the $O(\lambda^{4/3})$-term onwards.

6. Solution of the dual problem

We conclude by showing that the shadow price $\tilde{S}$ and the density process of the corresponding martingale measure are a solution of the dual problem from Definition 2.6.

Theorem 6.1. *Let Y and $\tilde{S}$ be defined as Lemma 4.5, and set*

$$Z = \mathscr{E}\left(\int_0^\cdot -\frac{\tilde{\mu}(Y_t)}{\tilde{\sigma}(Y_t)}dW_t\right).$$

Then

$$u_{(1-\lambda)S,S}(\eta_B,\eta_S) = v_{(1-\lambda)S,S}(\eta_B,\eta_S),$$

i.e., there is no duality gap in the market with bid-ask process $((1-\lambda)S,S)$, and $(\tilde{S},Z)$ is a dual optimizer.

Proof. First notice that for any continuous semimartingale $\bar{S}$ with values in the bid-ask spread $[(1-\lambda)S,S]$, it follows as in the proof [14, Theorem 4.6] that

$$u_{(1-\lambda)S,S}(\eta_B,\eta_S) \le u_{\bar{S},\bar{S}}(\eta_B,\eta_S).$$

Together with, e.g., [7, Lemma 2.3], this implies

$$u_{(1-\lambda)S,S}(\eta_B,\eta_S) \le v_{(1-\lambda)S,S}(\eta_B,\eta_S).$$

For the shadow price $\tilde{S}$, we have $v_{(1-\lambda)S,S}(\eta_B,\eta_S) \leq v_{\tilde{S},\tilde{S}}(\eta_B,\eta_S)$ by definition and $u_{\tilde{S},\tilde{S}}(\eta_B,\eta_S) = u_{(1-\lambda)S,S}(\eta_B,\eta_S)$ by Theorem 4.6. It therefore remains to show

$$v_{\tilde{S},\tilde{S}}(\eta_B,\eta_S) = u_{\tilde{S},\tilde{S}}(\eta_B,\eta_S),$$

i.e., that there is no duality gap in the frictionless market with price process $\tilde{S}$. As the frictionless market with price process $\tilde{S}$ is standard and complete in the sense of [15, Definition 1.7.3], this follows from [15, Theorem 3.9.12] and we are done. □

References

[1] Campi, L., and Schachermayer, W. A super-replication theorem in Kabanov's model of transaction costs. *Finance Stoch. 10*, 4 (2006), 579–596.

[2] Constantinides, G. Capital market equilibrium with transaction costs. *J. Polit. Econ.*, 94 (1986), 842–862.

[3] Cvitanić, J., and Karatzas, I. Hedging and portfolio optimization under transaction costs: a martingale approach. *Math. Finance 6*, 2 (1996), 133–165.

[4] Davis, M. H. A., and Norman, A. R. Portfolio selection with transaction costs. *Math. Oper. Res. 15*, 4 (1990), 676–713.

[5] Dumas, B., and Luciano, E. An exact solution to a dynamic portfolio choice problem under transaction costs. *J. Finance 46*, 2 (1991), 577–595.

[6] Gerhold, S., Muhle-Karbe, J., and Schachermayer, W. The dual optimizer for the growth-optimal portfolio under transaction costs. *Finance Stoch.* (2011). To appear.

[7] Goll, T., and Kallsen, J. Optimal portfolios for logarithmic utility. *Stochastic Process. Appl. 89*, 1 (2000), 31–48.

[8] Gould, H. W. Coefficient identities for powers of Taylor and Dirichlet series. *Amer. Math. Monthly 81* (1974), 3–14.

[9] Guasoni, P., Rásonyi, M., and Schachermayer, W. Consistent price systems and face-lifting pricing under transaction costs. *Ann. Appl. Probab. 18*, 2 (2008), 491–520.

[10] Gunning, R. C., and Rossi, H. *Analytic functions of several complex variables*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1965.

[11] Ilyashenko, Y., and Yakovenko, S. *Lectures on analytic differential equations*, vol. 86 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2008.

[12] Janeček, K., and Shreve, S. E. Asymptotic analysis for optimal investment and consumption with transaction costs. *Finance Stoch. 8*, 2 (2004), 181–206.

[13] Jouini, E., and Kallal, H. Martingales and arbitrage in securities markets with transaction costs. *J. Econom. Theory 66*, 1 (1995), 178–197.

[14] Kallsen, J., and Muhle-Karbe, J. On using shadow prices in portfolio optimization with transaction costs. *Ann. Appl. Probab. 20*, 4 (2010), 1341–1358.

[15] Karatzas, I., and Shreve, S. E. *Methods of mathematical finance*, vol. 39 of *Applications of Mathematics*. Springer-Verlag, New York, 1998.

[16] Kühn, C., and Stroh, M. Optimal portfolios of a small investor in a limit order market: a shadow price approach. *Math. Financ. Econ. 3*, 2 (2010), 45–72.

[17] Magill, M. J. P., and Constantinides, G. M. Portfolio selection with transactions costs. *J. Econom. Theory 13*, 2 (1976), 245–263.

[18] Merton, R. C. Lifetime portfolio selection under uncertainty: the continuous-time case. *Rev. Econ. Statist 51*, 3 (1969), 247–257.

[19] Merton, R. C. Optimum consumption and portfolio rules in a continuous-time model. *J. Econom. Theory 3*, 4 (1971), 373–413.

[20] Merton, R. C. *Continuous Time Finance*. Blackwell, Malden, 1990.

[21] Rogers, L. C. G. Why is the effect of proportional transaction costs $O(\delta^{2/3})$? In *Mathematics of finance*, vol. 351 of *Contemp. Math.* Amer. Math. Soc., Providence, RI, 2004, pp. 303–308.

[22] Schachermayer, W. The fundamental theorem of asset pricing under proportional transaction costs in finite discrete time. *Math. Finance 14*, 1 (2004), 19–48.

[23] Shreve, S. E., and Soner, H. M. Optimal investment and consumption with transaction costs. *Ann. Appl. Probab. 4*, 3 (1994), 609–692.

[24] Skorokhod, A. Stochastic equations for diffusion processes in a bounded region. *Theory Probab. Appl. 6* (1961), 264–274.
[25] Taksar, M., Klass, M. J., and Assaf, D. A diffusion model for optimal portfolio selection in the presence of brokerage fees. *Math. Oper. Res. 13*, 2 (1988), 277–294.
[26] Whalley, A. E., and Wilmott, P. An asymptotic analysis of an optimal hedging model for option pricing with transaction costs. *Math. Finance 7*, 3 (1997), 307–324.

Technische Universität Wien, Institut für Wirtschaftsmathematik
Wiedner Hauptstrasse 8-10, A-1040 Wien, Austria
E-mail address: sgerhold@fam.tuwien.ac.at

ETH Zürich, Departement für Mathematik
Rämistrasse 101, CH-8092, Zürich, Switzerland
E-mail address: johannes.muhle-karbe@math.ethz.ch

Universität Wien, Fakultät für Mathematik
Nordbergstrasse 15, A-1090 Wien, Austria
E-mail address: walter.schachermayer@univie.ac.at