3 The Growth-Optimal Portfolio in the Black-Scholes Model

In this section we follow the lines of [34] and analyze the dual optimizer in a Black-Scholes model under transaction costs $\lambda \geqslant 0$. The task is to maximize the expected return (or growth) of a portfolio. This is tantamount to consider utility maximization with respect to logarithmic utility $U(x) = \log(x)$ of terminal wealth at time T,

$$(P_x) \qquad \mathbb{E}\left[\log(V_T)\right] \to \max!, \qquad V_T \in x + C^{\lambda}. \tag{68}$$

Our emphasis will be on the limiting behavior for $T \to \infty$.

We take as stock price process $S = (S_t)_{t\geqslant 0}$ the Black-Scholes model

$$S_t = S_0 \exp\left[\sigma W_t + (\mu - \tfrac{\sigma^2}{2})t\right], \tag{69}$$

where $\sigma > 0$ and $\mu \geqslant 0$ are fixed constants.

To keep the notation light, the bond price process will again be assumed to be $B_t \equiv 1$. We remark that the case $B_t = \exp(rt)$ can rather trivially be reduced to the present one, simply by passing to discounted terms.

3.1 The frictionless case

We first recall the situation without transaction costs. This topic is well-known and goes back to the seminal work of R. Merton [71]. For later use we formulate the result in a slightly more general setting: we assume that the volatility σ and the drift μ are arbitrary predictable processes.

We fix the horizon T and assume that $W = (W_t)_{0\leqslant t\leqslant T}$ is a Brownian motion based on $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leqslant t\leqslant T}, \mathbb{P})$ where $(\mathcal{F}_t)_{0\leqslant t\leqslant T}$ is the (saturated) filtration generated by W.

Theorem 3.1. *(compare [71]): Suppose that the* $]0,\infty[$*-valued stock price process* $S = (S_t)_{0\leqslant t\leqslant T}$ *satisfies the stochastic differential equation*

$$\tfrac{dS_t}{S_t} = \mu_t dt + \sigma_t \, dW_t, \qquad 0 \leqslant t \leqslant T,$$

where $(\mu_t)_{0\leqslant t\leqslant T}$ *and* $(\sigma_t)_{0\leqslant t\leqslant T}$ *are predictable, real-valued processes such that*

$$\mathbb{E}\left[\int_0^T \frac{\mu_t^2}{\sigma_t^2} dt\right] < \infty. \tag{70}$$

Define the growth optimal process $\hat{V} = (\hat{V}_t)_{0\leqslant t\leqslant T}$, *starting at* $\hat{V}_0 = 1$, *by*

$$\frac{d\hat{V}_t}{\hat{V}_t} = \hat{\pi}_t \cdot \frac{dS_t}{S_t}, \tag{71}$$

where $\hat{\pi}_t$ *equals the mean variance ratio*

$$\hat{\pi}_t = \frac{\mu_t}{\sigma_t^2}. \tag{72}$$

Then $\hat{V}$ *is a well-defined* $]0,\infty[$*-valued process satisfying*

$$\hat{V}_t = \exp\left[\int_0^t \frac{\mu_s}{\sigma_s} dW_s + \int_0^t \frac{\mu_s^2}{2\sigma_s^2} ds\right], \qquad 0 \leqslant t \leqslant T. \tag{73}$$

We then have

$$\mathbb{E}\left[\log(\hat{V}_T)\right] = \mathbb{E}\Big[\int_0^T \frac{\mu_t^2}{2\sigma_t^2} dt\Big]. \tag{74}$$

If $(\pi_t)_{0\leqslant t\leqslant T}$ *is any competing strategy in* (71), *i.e. an* $\mathbb{R}$*-valued, predictable process such that*

$$\mathbb{E}\left[\int_0^T \pi_t^2\sigma_t^2 dt\right] < \infty, \text{ and } \int_0^T |\pi_t\mu_t| dt < \infty, \text{ a.s.}, \tag{75}$$

the stochastic differential equation

$$\frac{dV_t}{V_t} = \pi_t \frac{dS_t}{S_t} \tag{76}$$

well-defines a $]0,\infty[$*-valued process*

$$V_t = \exp\left[\int_0^t \pi_s\sigma_s dW_s + \int_0^t \left(\pi_s\mu_s - \frac{\pi_s^2\sigma_s^2}{2}\right) ds\right], \tag{77}$$

for which we obtain

$$\mathbb{E}\left[\log(V_T)\right] \leqslant \mathbb{E}\left[\log(\hat{V}_T)\right],$$

and, more generally, for stopping times $0 \leqslant \varrho \leqslant \tau \leqslant T$

$$\mathbb{E}\left[\log(\tfrac{V_\tau}{V_\varrho})\right] \leqslant \mathbb{E}\left[\log(\tfrac{\hat{V}_\tau}{\hat{V}_\varrho})\right].$$

Proof: If a strategy $(\pi_t)_{0\leqslant t\leqslant T}$ satisfies (75) we get from Itô's formula and (69) that (77) is the solution to (76) with initial value $V_0 = 1$. Passing to $\hat{\pi}$ defined in (72), the assertion (74) is rather obvious

$$\mathbb{E}\left[\log(\hat{V}_T)\right] = \mathbb{E}\left[\int_0^T \tfrac{\mu_t}{\sigma_t} dW_t + \int_0^T \tfrac{\mu_t^2}{2\sigma_t^2} dt\right] = \mathbb{E}\left[\int_0^T \tfrac{\mu_t^2}{2\sigma_t^2} dt\right]$$

as $(\int_0^t \frac{\mu_s}{\sigma_s} dW_s)_{0\leqslant t\leqslant T}$ is a martingale bounded in $L^2(\mathbb{P})$ by (70).

If $\pi = (\pi_t)_{0\leqslant t\leqslant T}$ is any competing strategy verfying (75), we again obatin

$$\mathbb{E}\left[\log(V_T)\right] = \mathbb{E}\left[\int_0^T \pi_t\sigma_t dW_t + \int_0^T \left(\pi_t\mu_t - \frac{\pi_t^2\sigma_t^2}{2}\right) dt\right] = \mathbb{E}\left[\int_0^T \left(\pi_t\mu_t - \frac{\pi_t^2\sigma_t^2}{2}\right) dt\right].$$

It is obvious that, for fixed $0 \leqslant t \leqslant T$ and $\omega \in \Omega$, the function

$$\pi \to \pi\mu_t(\omega) - \frac{\pi^2\sigma_t^2(\omega)}{2}, \qquad \pi \in \mathbb{R},$$

attains its unique maximum at $\hat{\pi}_t(\omega) = \frac{\mu_t(\omega)}{\sigma_t^2(\omega)}$ so that

$$\mathbb{E}\left[\log(V_T)\right] \leqslant \mathbb{E}\left[\int_0^T \left(\hat{\pi}_t\mu_t - \frac{\hat{\pi}_t^2\sigma_t^2}{2}\right) dt\right] = \mathbb{E}\left[\int_0^T \frac{\mu_t^2}{2\sigma_t^2} dt\right] = \mathbb{E}\left[\log(\hat{V}_T)\right].$$

More generally, for stopping times $0 \leqslant \varrho \leqslant \tau \leqslant T$, we obtain

$$\mathbb{E}\left[\log\left(\frac{V_\tau}{V_\varrho}\right)\right] = \mathbb{E}\left[\int_\varrho^\tau \left(\pi_t\mu_t - \frac{\pi_t^2\sigma_t^2}{2}\right) dt\right] \leqslant \mathbb{E}\left[\int_\varrho^\tau \left(\hat{\pi}_t\mu_t - \frac{\hat{\pi}_t^2\sigma_t^2}{2}\right) dt\right] = \mathbb{E}\left[\log\left(\frac{\hat{V}_\tau}{\hat{V}_\varrho}\right)\right].$$

■

3.2 Passing to transaction costs: some heuristics

Before we pass to a precise formulation of the utility maximization problem for the log-utility maximizer (see Definition 3.9 below) we want to develop the heuristics to find the shadow price process $(\tilde{S}_t)_{t\geqslant 0}$ for the utility maximization problem of optimizing the expected growth of a portfolio. We make two heroic assumptions. In fact, we are allowed to make all kind of heuristic assumptions and bold guesses, as we shall finally pass to verification theorems to justify them.

Assumption 3.2. *When the shadow price $(\tilde{S}_t)_{t\geqslant 0}$ ranges in the interior $](1-\lambda)S_t, S_t[$ of the bid-ask interval $[(1-\lambda)S_t, S_t]$ then the process $\tilde{S}_t$ is a deterministic function of S_t*

$$\tilde{S}_t = g_c(S_t). \tag{78}$$

More precisely, we suppose that there is a family of (deterministic, smooth) functions $g_c(\cdot)$, depending on a real parameter c, such that, whenever we have random times $\varrho \leqslant \tau$ such that $\tilde{S}_t \in](1-\lambda)S_t, S_t[$, for all $t \in]\!]\varrho, \tau[\![$, then there is a fixed parameter c (depending on the interval $]\!]\varrho, \tau[\![$) such that

$$\tilde{S}_t = g_c(S_t), \qquad \varrho \leqslant t \leqslant \tau.$$

The point is that the parameter c does not change while $\tilde{S}_t$ ranges in the interior $](1-\lambda)S_t, S_t[$ of the bid-ask interval. Only when $\tilde{S}_t$ equals $(1-\lambda)S_t$ or S_t we shall allow the parameter c to vary.

Assumption 3.3. *A log-utility agent, who can invest in* a frictionless way *(i.e. without paying transaction costs) in the market $\tilde{S}$* does not want to change her positions in stock and bond *as long as $\tilde{S}_t$ ranges in the interior $](1-\lambda)S_t, S_t[$ of the bid-ask interval.*

Assumption 3.3 is, of course, motivated by the results on the shadow price process in Section 2 (Def. 2.7).

Here are two consequences of the above assumptions. Suppose that $\tilde{S}_t$ satisfies

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = \tilde{\mu}_t dt + \tilde{\sigma}_t dW_t, \tag{79}$$

where $(\tilde{\mu}_t)_{t\geqslant 0}$ and $(\tilde{\sigma}_t)_{t\geqslant 0}$ are general predictable processes which we eventually want to determine. Applying Itô to (78) and dropping the subscript c of

g_c for the moment (and supposing that g is sufficiently smooth), we obtain $d(g(S_t)) = g'(S_t)dS_t + \frac{g''(S_t)}{2}\, d\langle S\rangle_t$, or, equivalently

$$\frac{d\tilde{S}_t}{\tilde{S}_t} = \frac{g'(S_t)}{g(S_t)}\, dS_t + \frac{g''(S_t)}{2g(S_t)}\, d\langle S\rangle_t.$$

Inserting (69) we obtain in (79) above (compare [34])

$$\tilde{\sigma}_t = \frac{\sigma g'(S_t)S_t}{g(S_t)} \tag{80}$$

$$\tilde{\mu}_t = \frac{\mu g'(S_t)S_t + \frac{\sigma^2}{2} g''(S_t)S_t^2}{g(S_t)} \tag{81}$$

and in particular the relation

$$\frac{\tilde{\mu}_t}{\tilde{\sigma}_t^2} = \frac{g(S_t)[\mu g'(S_t)S_t + \frac{\sigma^2}{2} g''(S_t)S_t^2]}{\sigma^2 g'(S_t)^2 S_t^2}. \tag{82}$$

On the other hand, it follows from Theorem 3.1 that the optimal proportion $\tilde{\pi}$ of the investment $\varphi^1\tilde{S}$ into stock to total wealth $\varphi^0 + \varphi^1\tilde{S}$ for the log-utility optimizer in the frictionless market $\tilde{S}$ is given by

$$\tilde{\pi}_t = \frac{\varphi_t^1\tilde{S}_t}{\varphi_t^0 + \varphi_t^1\tilde{S}_t} = \frac{g(S_t)}{c + g(S_t)}, \tag{83}$$

where

$$c := \frac{\varphi_t^0}{\varphi_t^1} \tag{84}$$

is the ratio of positions φ_t^0 and φ_t^1 in bond and stock respectively. Assumption 3.3 implies that φ_t^0 and φ_t^1, and therefore also the parameter c, should remain constant when $\tilde{S}$ ranges in the interior $](1-\lambda)S_t, S_t[$ of the bid-ask spread.

We have assembled all the ingredients to yield a unifying equation: on the one hand side, the ratio $\tilde{\pi}_t$ of the value of the investment in stock and total wealth (both evaluated by using the shadow price $\tilde{S}$) is given by formula (83). On the other hand, by formula (72) in Theorem 3.1 and Assumption 3.3 we must have $\tilde{\pi}_t = \frac{\tilde{\mu}_t}{\tilde{\sigma}_t^2}$ and the latter ratio is given by (82). Hence

$$\tilde{\pi}_t = \frac{g(S_t)}{c + g(S_t)} = \frac{g(S_t)[\mu g'(S_t)S_t + \frac{\sigma^2}{2} g''(S_t)S_t^2]}{\sigma^2 g'(S_t)^2 S_t^2}.$$

Rearranging this equation and substituting S_t by the variable $s \in \mathbb{R}_+$, we arrive at the ODE

$$g''(s) = \frac{2g'(s)^2}{c + g(s)} - \frac{2\mu g'(s)}{\sigma^2 s}, \qquad s > 0. \tag{85}$$

Somewhat surprisingly this ODE admits a closed form solution (compare, however, Section 3.9 below for a good reason why we actually find a closed form solution). Before spelling out this solution let us pass to a (heuristic) discussion of the initial conditions of the ODE (85). Fix $t_0 \geqslant 0$ and suppose that we have $S_{t_0} = 1$ which is just a matter of normalization. More importantly, suppose also that $\tilde{S}_{t_0} = S_{t_0} = 1$. The economic interpretation is that the economic agent was just buying stock at time t_0 which forces the shadow price $\tilde{S}_{t_0}$ to equal the ask price S_{t_0}. We also suppose (very heuristically!) that $(S_t)_{t\geqslant 0}$ starts a positive excursion at time t_0, i.e. $S_t > S_{t_0}$ for $t > t_0$ such that $t - t_0$ is sufficiently small.

We then are led to the initial conditions for (85)

$$g(1) = 1, \quad g'(1) = 1. \tag{86}$$

The second equation is a "smooth pasting condition" requiring that S_t and $\tilde{S}_t = g(S_t)$ match of first order around $t = t_0$. The necessity of this condition is intuitively rather clear and will become obvious in subsection 3.7 below.

We write $\theta = \frac{\mu}{\sigma^2}$ as (85) only depends on this ratio. As Mathematica tells us, the general form of the solution to (85) satisfying the initial conditions (86) then is given by

$$g(s) = g_c(s) = \frac{-cs + (2\theta - 1 + 2c\theta)s^{2\theta}}{s - (2 - 2\theta - c(2\theta - 1))s^{2\theta}} \tag{87}$$

unless $\theta = \frac{1}{2}$, which is a special case (see (88) below) that can be treated analogously. The parameter c defined in (84) is still free in (87).

As regards the given mean-variance ratio $\theta = \frac{\mu}{\sigma^2} > 0$, we have to distinguish the regimes $\theta \in]0,1[$, $\theta = 1$, and $\theta > 1$. Let us start by discussing the singular case $\theta = 1$: in this case (see Theorem 3.1) the optimal solution in the frictionless market $S = (S_t)_{t\geqslant 0}$ defined in (69) is given by $\hat{\pi}_t \equiv 1$. Speaking economically, the utility maximizing agent, at time $t = 0$, invests all her wealth into stock and keeps this position unchanged until maturity T. In other words, no dynamic trading takes place in this special case, even without transaction costs. We therefore expect that this case will play a special (degenerate) role when we pass to transaction costs $\lambda > 0$.

The singular case $\theta = 1$ divides the regime $\theta \in]0,1[$ from the regime $\theta > 1$. In the former the log-utility maximizer holds positive investments in stock as well as in bond, while in the latter case she goes short in bond and invests more than her total wealth into stock. These well-known facts follow immediately from Theorem 3.1 in the frictionless case and we shall see in

Theorem 3.10 below that this basic feature still holds true in the presence of transaction costs, at least for $\lambda > 0$ sufficiently small.

The mathematical analysis reveals that the case $\theta = \frac{1}{2}$ also plays a special role (apart from the singular case $\theta = 1$): in this case the general solution to the ODE (85) under initial conditions (86) involves logarithmic terms rather than powers:

$$g(s) = g_c(s) = \frac{c + 1 + c\ \log(s)}{c + 1 - \log(s)}. \quad (88)$$

But this solution is only special from a mathematical point of view while, from an economic point of view, this case is not special at all and we shall see that the solution (88) nicely interpolates the solution (87), for $\theta \to \frac{1}{2}$.

We now pass to the elementary, but tedious, discussion of the qualitative properties of the functions $g_c(\cdot)$ in (87) and (88) respectively. As this discussion amounts - at least in principle - to an involved version of a high school exercise, we only resume the results and refer for proofs to [34, Appendix A].

3.3 The case $0 < \theta < 1$

In this case we consider the function $g(s) = g_c(s)$ given by (87) and (88) respectively, on the right hand side of $s = 1$, i.e. on the domain $s \in [1, \infty[$. Fix the parameter c in $]\frac{1-\theta}{\theta}, \infty[$ for $\theta \in]0, \frac{1}{2}]$ (resp. in $]\frac{1-\theta}{\theta}, \frac{1-\theta}{\theta - \frac{1}{2}}[$ for $\theta \in]\frac{1}{2}, 1[$). Plugging $s = 1$ into the ODE (85) we observe that the above domains were chosen in such a way to have $g_c''(1) < 0$. Hence for fixed $c \in]\frac{1-\theta}{\theta}, \infty[$ (resp. $c \in]\frac{1-\theta}{\theta}, \frac{1-\theta}{\theta - \frac{1}{2}}[$ for $\theta \in]\frac{1}{2}, 1[$) the function $g_c(\cdot)$ is strictly concave in a neighbourhood of $s = 1$ so that from (86) we obtain

$$g_c(s) < s,$$

for $s \neq 1$ sufficiently close to $s = 1$.

Figure 3 is a picture of the qualitative features of the function $g_c(\cdot)$ on $s \in [1, \hat{s}[$. The point $\hat{s} > 1$ is the pole of $g_c(\cdot)$ where the denominator in (87) (resp. (88)) vanishes.

The function g_c is strictly increasing on $[1, \hat{s}[$; it is concave in a neighborhood of $s = 1$, then has a unique inflection point in $]1, \hat{s}[$, and eventually is convex between the inflection point and the pole $\hat{s}$.

We also observe that, for $\frac{1-\theta}{\theta} < c_1 < c_2$ we have $g_{c_1}(s) > g_{c_2}(s)$, for $s \in [1, \hat{s}[$, where $\hat{s}$ is the pole of the function g_{c_1} as displayed in Figure 4.

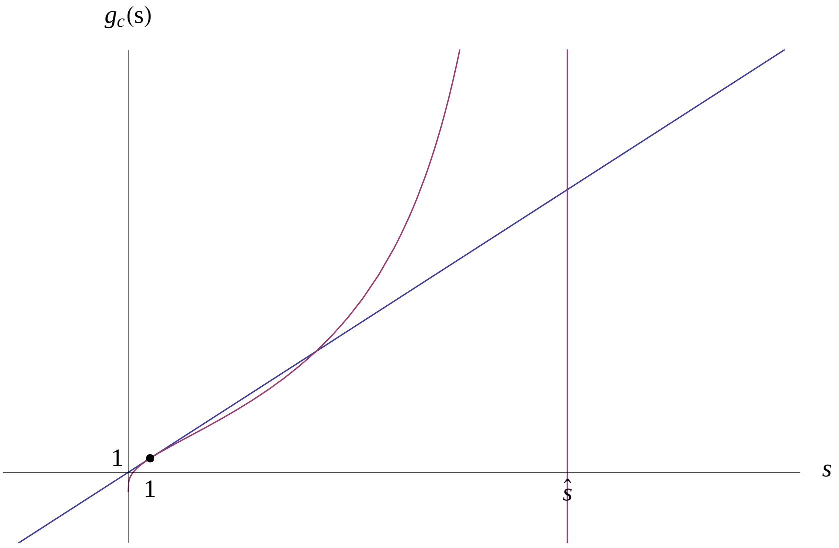

Figure 3: The function $g_c(s)$.

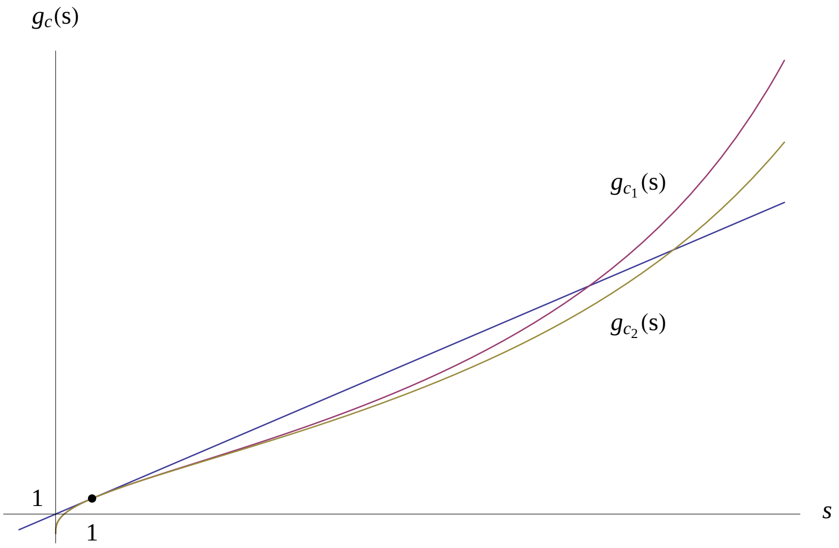

Figure 4: The functions $g_{c_1}(s)$ and $g_{c_2}(s)$, for $c_1 < c_2$.

We still have to complement the boundary conditions (86) for the ODE (85) at the other endpoint, corresponding to the “selling boundary”: we want to find a point $\bar{s} = \bar{s}(c) \in\,]1, \hat{s}[$ and $0 < \lambda < 1$ such that

$$g_c(\bar{s}) = (1-\lambda)\bar{s}, \qquad\qquad g_c'(\bar{s}) = (1-\lambda). \tag{89}$$

Geometrically this task corresponds to drawing the unique line through the origin which tangentially touches the graph of $g_c(\cdot)$. See Figure 5.

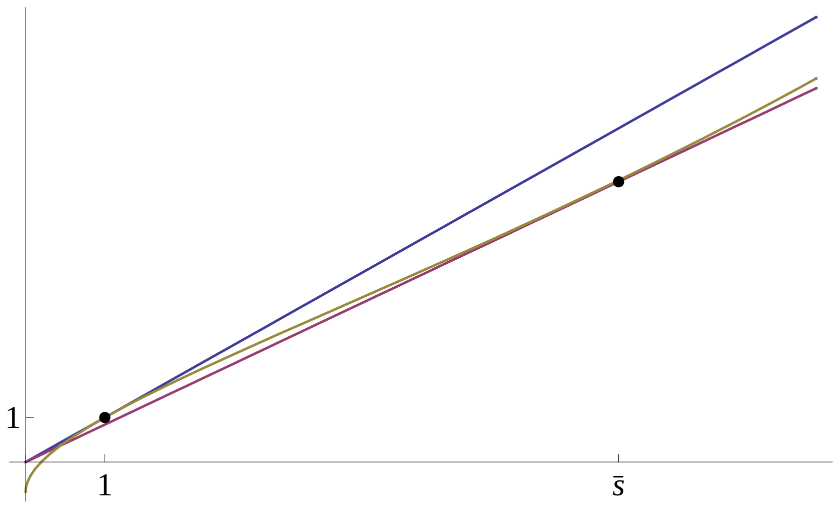

Figure 5: Smooth pasting conditions for the function g.

If we have found this tangent and the touching point $\bar{s}$, then (89) holds true, where $(1-\lambda)$ is the slope of the tangent.

In fact, for fixed $c \in\,]\frac{1-\theta}{\theta}, \infty[$ and $\theta \in]0, \frac{1}{2}]$ (resp. $c \in\,]\frac{1-\theta}{\theta}, \frac{1-\theta}{\theta-\frac{1}{2}}[$ and $\theta \in]\frac{1}{2}, 1[$) one may explicitly solve the two equations (89) in the two variables λ and $\bar{s}$ by simply plugging in formula (87) to obtain, for $\frac{1-\theta}{\theta} < c < \infty$,

$$\bar{s} = \bar{s}(c) \;=\; \left(\frac{c}{(2\theta - 1 + 2c\theta)(2 - 2\theta - c(2\theta-1))}\right)^{1/(2\theta-1)}, \tag{90}$$

$$\lambda = \lambda(c) \;=\; \frac{(1-2(c+1)\theta)\bar{s}(c)^{2\theta} + c\bar{s}(c)}{\bar{s}(c)\left((2(c+1)\theta - c - 2)\bar{s}(c)^{2\theta} + \bar{s}(c)\right)} + 1, \tag{91}$$

$$g(\bar{s}) = \tag{92}$$

$$\frac{(2(c+1)\theta-1)\left(\left(-\frac{(2(c+1)\theta-1)(2(c+1)\theta-c-2)}{c}\right)^{\frac{1}{1-2\theta}}\right)^{2\theta} - c\left(-\frac{(2(c+1)\theta-1)(2(c+1)\theta-c-2)}{c}\right)^{\frac{1}{1-2\theta}}}{(2(c+1)\theta-c-2)\left(\left(-\frac{(2(c+1)\theta-1)(2(c+1)\theta-c-2)}{c}\right)^{\frac{1}{1-2\theta}}\right)^{2\theta} + \left(-\frac{(2(c+1)\theta-1)(2(c+1)\theta-c-2)}{c}\right)^{\frac{1}{1-2\theta}}}.$$

In the special case $\theta = \frac{1}{2}$, where $\frac{1-\theta}{\theta} = 1$, we obtain the somewhat simpler formulae

$$\bar{s} = \bar{s}(c) = \exp\left(\frac{c^2 - 1}{c}\right), \qquad 1 < c < \infty, \quad (93)$$

$$\lambda = \lambda(c) = 1 - c^2 \exp\left(\frac{1 - c^2}{c}\right), \qquad 1 < c < \infty, \quad (94)$$

$$g(\bar{s}) = g_c(\bar{s}(c)) = c^2, \qquad 1 < c < \infty. \quad (95)$$

We summarize what we have found so far.

Proposition 3.4. *Fix $\theta \in]0, 1[$ and $c \in]\frac{1-\theta}{\theta}, \infty[$ (resp. $c \in]\frac{1-\theta}{\theta}, \frac{1-\theta}{\theta - \frac{1}{2}}[$ if $\theta \in]\frac{1}{2}, 1[$). Then the function $g(s) = g_c(s)$ defined in* (87) *(resp.* (88)*) is strictly increasing in $[1, \bar{s}]$, where $\bar{s} = \bar{s}(c)$ is defined in* (90) *(resp.* (93)*). In addition, g satisfies the boundary conditions*

$$g_c(1) = 1, \qquad g_c'(1) = 1,$$
$$g_c(\bar{s}) = (1 - \lambda)\bar{s}, \qquad g_c'(\bar{s}) = 1 - \lambda,$$

where λ is given by (91) *(resp.* (94)*).*

Proof: The energetic reader may verify the above assertions by simply calculating all the above expressions and discussing the function g_c'. ■

The drawback of the above proposition is that c is the free variable parameterizing the solution. The transaction costs $\lambda = \lambda(c)$ in (91) (resp. (94)) are a function of c. Our original problem, however is stated the other way round: the level $0 < \lambda < 1$ of transaction costs is given and c as well as $\bar{s} = \bar{s}(c)$ and the function $g = g_c$ depend on λ. In other words, we have to invert the formulae (91) and (94). Unfortunately, when we shall do this final step, we will have to leave the pleasant case of closed form solutions which we have luckily encountered so far. We shall only be able to determine the inverse function of (91) (resp. (94)) locally around $\lambda = 0$ as a fractional Taylor series in λ (see (97) below). As this Taylor series only converges in some neighborhood of $\lambda = 0$, from now on, every assertion has to be preceded by the caveat "for $\lambda > 0$ sufficiently small". Hence we are interested in the behavior of the function $\lambda = \lambda(c)$ in (91) (resp. (94)) when c is in a neighborhood of the left limit $\frac{1-\theta}{\theta}$ of its domain: this corresponds to λ being in a neighborhood of zero.

In order to keep the calculations simple we focus on the special case $\theta = \frac{1}{2}$. The arguments carry over to the case of general $0 < \theta < 1$, at the expense of somewhat longer formulae (compare [34]).

Differentiating $\lambda(c)$ in (94) with respect to c we obtain

$$\lambda'(c) = (c-1)^2 \exp\left(\tfrac{1-c^2}{c}\right),$$
$$\lambda''(c) = (c-1)\exp(\tfrac{1-c^2}{c})\tfrac{-3c^3+5c^2+c-1}{c^2},$$
$$\lambda'''(c) = \frac{1+c^2(3+c(-6+(-3+c)^2c))}{c^4}\exp\left(\frac{1}{c}-c\right)$$

so that $\lambda(1) = \lambda'(1) = \lambda''(1) = 0$ while $\lambda'''(1) = 2 \neq 0$. Therefore the Taylor expansion of the analytic function $\lambda(c)$ around $c = 1$ starts as

$$\lambda(c) = \tfrac{1}{3}(c-1)^3 + O(c-1)^4.$$

This implies that the function $c \mapsto \lambda(c)$ given in (94) $\lambda = 0$ is locally invertible around $c = 1$ and that the inverse function $\lambda \mapsto c(\lambda)$ has a fractional Taylor expansion in terms of powers of $\lambda^{1/3}$ around $\lambda = 0$, with leading term

$$c(\lambda) = 1 + 3^{1/3}\lambda^{1/3} + O(\lambda^{2/3}). \tag{96}$$

As shown in [34] one may algorithmically determine all the coefficients in the above fractional Taylor expansion (96) of the function $\lambda \mapsto c(\lambda)$. This not only works for the specially simple case $\theta = \frac{1}{2}$ considered above, but for all $\theta \in \,]0,1[$ and the coefficients are explicit functions of θ, which turn out to be fractional powers of certain rational functions of θ (see Proposition 3.5 below as well as Proposition 6.1 of [34] for the details).

Once we have expanded the parameter c as a function of λ around $\lambda = 0$ we can, for $c = c(\lambda)$, also plug this expansion into all the other quantities depending on c, e.g. $\bar{s} = \bar{s}(c)$ given in (90) (resp. (93)), to again obtain fractional Taylor expansions in λ. We resume our findings in the next proposition and refer to [34] for details and full proofs.

Proposition 3.5. *Fix $\theta \in]0,1[$. There are fractional Taylor series*

$$c(\lambda) = \frac{1-\theta}{\theta} + \frac{1-\theta}{2\theta}\left(\frac{6}{\theta(1-\theta)}\right)^{1/3}\lambda^{1/3} + \frac{(1-\theta)^2}{4\theta}\left(\frac{6}{\theta(1-\theta)}\right)^{2/3}\lambda^{2/3} + O(\lambda) \tag{97}$$

$$\bar{s}(\lambda) = 1 + \left(\frac{6}{\theta(1-\theta)}\right)^{1/3}\lambda^{1/3} + \frac{1}{2}\left(\frac{6}{\theta(1-\theta)}\right)^{2/3}\lambda^{2/3} + O(\lambda) \tag{98}$$

such that, for $\lambda \geqslant 0$ sufficiently small, the above series converge. The function $g(s) = g_{c(\lambda)}(s)$, defined on the interval $[1, \bar{s}(\lambda)]$ and given in (87) *(resp.*

(88)*), then satisfies the ODE* (85) *as well as the boundary conditions*

$$\begin{aligned} g(1) &= 1, & g'(1) &= 1, \\ g(\bar{s}(\lambda)) &= (1-\lambda)\bar{s}(\lambda), & g'(\bar{s}(\lambda)) &= (1-\lambda). \end{aligned}$$

■

3.4 Heuristic construction of the shadow price process

Fix $\theta \in \,]0,1[$ and $\lambda > 0$ as in the previous proposition. We shall continue to do some heuristics in this sub-section to motivate the sub-sequent formal definition. Define

$$\tilde{S}_t = g(S_t), \qquad t \geqslant 0, \tag{99}$$

where $g = g_{c(\lambda)}$ was defined in (87) and $c(\lambda)$ in (97).

Normalize S to satisfy $S_0 = 1$ so that also $\tilde{S}_0 = g(S_0) = 1$, and suppose (again heuristically!) that S starts a positive excursion at time $t = 0$, i.e. that $S_t > 1$ for $t > 0$ sufficiently small. In sub-section 3.2 the function g has been designed in such a way that the log-utility optimizer in the frictionless market $\tilde{S}$ keeps her holdings φ_t^0 and φ_t^1 constant, where the ratio $\frac{\varphi_t^0}{\varphi_t^1} = \frac{\varphi_t^0}{\varphi_t^1 \tilde{S}_0}$ equals the constant $c = c(\lambda)$ (in (97)).

But what happens if S_t hits the boundaries 1 or $\bar{s}$ of the interval $[1, \bar{s}]$? Say, at time $t_0 > 0$ we have for the first time after $t = 0$ that again we have $S_{t_0} = 1$. Consider the Brownian motion $W = (W_t)_{t \geqslant 0}$ during the infinitesimal interval $[t_0, t_0 + dt]$.

Interpreting, following a good tradition applied in physics, W as a random walk on an infinitesimal grid, we have (heuristically!) two possibilities for the increment of W : either $dW_{t_0} := W_{t_0+dt} - W_{t_0} = dt^{1/2}$ or $dW_{t_0} := W_{t_0+dt} - W_{t_0} = -dt^{1/2}$.

Let us start with the former case: we then have $dS_{t_0} = S_{t_0}(\mu\, dt + \sigma\, dt^{1/2})$ so that, continuing to define $\tilde{S}$ by (99)

$$\begin{aligned} d\tilde{S}_{t_0} := g(S_{t_0+dt}) - g(S_{t_0}) &= g'(S_{t_0})dS_{t_0} + \tfrac{1}{2}g''(S_{t_0})d\langle S\rangle_{t_0} \\ &= S_{t_0}(\mu\, dt + \sigma\, dt^{1/2}) + \frac{g''(1)}{2}S_{t_0}^2\sigma^2 dt \\ &= \sigma\, dt^{1/2} + \left(\mu + \frac{g''(1)}{2}\sigma^2\right) dt. \end{aligned} \tag{100}$$

Note that $g''(1) = \frac{2}{c+1} - 2\theta < 0$, as follows from (80).

The case $dW_{t_0} = -dt^{1/2}$ is different from the case $dW_{t_0} = +dt^{1/2}$: in this case we cannot blindly use definition (99) to find $\tilde{S}_{t_0+dt}$, as S_{t_0+dt} is (infinitesimally) outside the domain of definition $[1, \bar{s}]$ of g. In this case we move $\tilde{S}$ identically to S: in Fig. 4 this corresponds geometrically to the fact that $\tilde{S}$ decreases along the identity line. We then get

$$\begin{aligned} d\tilde{S}_{t_0} = dS_{t_0} &= S_{t_0}(\mu \; dt - \sigma \; dt^{1/2}) \\ &= -\sigma \; dt^{1/2} + \mu \; dt. \end{aligned} \tag{101}$$

When S_{t_0} thus has moved out of the domain $[1, \bar{s}]$ of g, the agent also has to rebalance the portfolio $(\varphi_t^0, \varphi_t^1)$ in order to keep the ratio of wealth in bond and wealth in stock

$$c = \frac{\varphi_{t_0}^0}{\varphi_{t_0}^1 \tilde{S}_{t_0}} = \frac{\varphi_{t_0+dt}^0}{\varphi_{t_0+dt}^1 \tilde{S}_{t_0+dt}} \tag{102}$$

constant. This is achieved by buying an infinitesimal amount (of order $dt^{1/2}$) of stock at ask price $S_{t_0} = \tilde{S}_{t_0} = 1$. In order for (102) to match with (101) we must have

$$d\varphi_{t_0}^1 = \varphi_{t_0}^1 \frac{c}{c+1} \; \sigma \; dt^{1/2}, \quad d\varphi_{t_0}^0 = -\varphi_{t_0}^0 \frac{1}{c+1} \; \sigma \; dt^{1/2} \tag{103}$$

as one easily checks by plugging (103) into (102) (neglecting terms of higher order than $dt^{1/2}$). Note in passing that $\tilde{S}_{t_0+dt} = S_{t_0+dt}$ also corresponds to the last fact that the agent is buying stock during the infinitesimal interval $[t_0, t_0 + dt]$.

We continue the discussion of the case $W_{t_0+dt} - W_{t_0} = -dt^{1/2}$ by passing to the next infinitesimal interval $[t_0 + dt, t_0 + 2dt]$: again we have to distinguish the case $W_{t_0+2dt} - W_{t_0+dt} = +dt^{1/2}$ and $W_{t_0+2dt} - W_{t_0+dt} = -dt^{1/2}$. Let us first consider the second case: we then continue to move $\tilde{S}$ in an identical way as S (compare (101)) and to keep buying stock at price S_{t_0+dt} which yields the same formula as in (103) , neglecting again terms of higher order than $dt^{1/2}$.

But what do we do if $W_{t_0+2dt} - W_{t_0+dt} = +dt^{1/2}$? The intuition is that we now move again into the no-trade region, where $\tilde{S}$ should depend on S in a functional way, similarly as in (99). This is indeed the case, *but the function g now has to be rescaled.* The domain of definition $[1, \bar{s}]$ has to be replaced by the interval $[m_t, m_t \bar{s}]$, where $(m_t)_{t \geqslant 0}$ denotes the (local) running minimum of the process $(S_t)_{t \geqslant 0}$: in our present infinitesimal reasoning (neglecting terms of higher order than $dt^{1/2}$) we have $m_{t_0+dt} = S_{t_0+dt} = 1 - \sigma \; dt^{1/2}$. If $(S_t)_{t \geqslant t_0+dt}$ starts a positive excursion at time $t_0 + dt$, which heuristically corresponds to $W_{t_0+2dt} - W_{t_0+dt} = +dt^{1/2}$, we define $\tilde{S}$ by

$$\tilde{S}_t = m_t \; g\left(\frac{S_t}{m_t}\right), \qquad t \geqslant t_0 + dt, \tag{104}$$

where $t \geqslant t_0 + dt$ is sufficiently small so that $(S_t)_{t \geqslant t_0+dt}$ remains above $m_{t_0+dt} = S_{t_0+dt} = 1 - \sigma \, dt^{1/2}$.

We have used the term m_t rather than $1 - \sigma dt^{1/2}$ in order to indicate that the previous formula not only holds true for the infinitesimal reasoning, but also for finite movements by considering the running minimum process $m_t = \inf_{0 \leqslant u \leqslant t} S_u$.

Summing up: during positive excursions of $(S_t)_{t \geqslant 0}$ we expect the process $(\tilde{S}_t)_{t \geqslant 0}$ to be defined by forumla (104), while at times t when $(S_t)_{t \geqslant 0}$ hits its running minimum $m_t = \min_{0 \leqslant u \leqslant t} S_u$ we simply let $\tilde{S}_t = S_t$ and buy stock similarly as in (103), following the movements of the running minimum $(m_t)_{t \geqslant 0}$.

The behavior of $\tilde{S}$ might remind of a reflected diffusion: by (104), we always have $\tilde{S}_t \leqslant S_t$, with equality happening when S_t equals its running minimum m_t. It is well-known that the set $\{t \in \mathbb{R}_+ : S_t = m_t\}$ is a Cantor-like subset of $\mathbb{R}_+$ of Lebesgue measure, related to "local time". There is, however an important difference between the present situation and, say, reflected Brownian motion $(|W_t|)_{t \geqslant 0}$: we shall prove below that $(\tilde{S}_t)_{t \geqslant 0}$ is a *diffusion*, i.e. its semi-martingale characteristics are absolutely continuous with respect to Lebesgue measure. In other words, the process $(\tilde{S}_t)_{t \geqslant 0}$ does not involve a "local time component". The reason for this remarkable feature of $\tilde{S}$ is the smooth pasting condition $g'(1) = 1$ in (87). This condition yielded in the above calculations that the leading terms of the differentials (100) and (101) are - up to the sign - identical, namely $\sigma \, dt^{1/2}$ and $-\sigma \, dt^{1/2}$. In other words, when $m_t = S_t = \tilde{S}_t$ so that the movement of $\tilde{S}$ is given by the regime (100) or (101), the effect of order $dt^{1/2}$ on the movement of $\tilde{S}_t$ is given by $\sigma \, dW_t$ as the leading terms in (100) and (101) are symmetric. This distinguishes the behavior of the process $\tilde{S}$ from, e.g., reflected Brownian motion where this relation fails to be symmetric when reflection takes place.

A closer look at the differentials (100) and (101) reveals that the terms of order dt do *not coincide* any more. However, this will do no harm, as the set of time instances t where $S_t = \tilde{S}_t$, i.e. S_t equals its running minimum m_t, only is a set of Lebesgue measure zero. Integrating quantities of order dt over such a set will have no effect.

The fact that the terms of order dt do not coincide in (100) and (101) corresponds to the fact that the extended function $G : [0, \bar{s}] \to [0, (1 - \lambda)\bar{s}$

$$G(s) = \begin{cases} s, & \text{for } 0 \leqslant s \leqslant 1 \\ g(s), & \text{for } 1 \leqslant s \leqslant \bar{s} \end{cases}$$

is once, but *not twice* differentiable: the second derivative is discontinuous at the point $s = 1$ (with finite left and right limits). It is well known that such an isolated discontinuity of the second derivate does not restrict the applicability of Itô's lemma, which is the more formal version of the above heuristics.

Here is another aspect to be heuristically discussed before we turn to the mathematically precise formulation (Theorem 3.6 below) of the present theme. So far we have only dealt with the case when the process $\tilde{S}_t$ equals the ask price S_t or makes some (small) excursion away from it. We still have to discuss the behavior of $\tilde{S}$ when it makes a "large" excursion, so that $\tilde{S}_t$ hits the bid price $(1-\lambda)S_t$. In this case an analogous phenomenon happens, with signs reversed.

To fix ideas, suppose again (heuristically) that the process $(S_t)_{t\geqslant 0}$ starts a positive excursion at $S_0 = 1$ and hits the level $\bar{s} > 1$ at some time $t_1 > 0$. We then have, in accordance with (99),

$$\tilde{S}_t = g(S_t), \qquad 0 \leqslant t \leqslant t_1, \tag{105}$$

and $\tilde{S}_{t_1} = g(\bar{s}) = (1-\lambda)S_{t_1}$, i.e. $\tilde{S}_t$ hits the bid price $(1-\lambda)S_t$ at time $t = t_1$. What happens now? Again we distinguish the cases $dW_{t_1} = W_{t_1+dt} - W_{t_1} = \pm dt^{1/2}$. If $dW_{t_1} = -dt^{1/2}$, we turn back into the no-trade region: we continue to define $\tilde{S}$ via (105) also at time $t_1 + dt$. If, however $dW_{t_1} = +dt^{1/2}$ we define

$$\tilde{S}_{t_1+dt} = (1-\lambda)S_{t_1+dt},$$

i.e., the relation between $\tilde{S}$ and S is given by the straight line through the origin with slope $1-\lambda$ (see Figure 5). We then sell stock at the bid price $\tilde{S}_{t_1} = (1-\lambda)S_{t_1}$ in a similar way as in (103), but now with the signs of $d\varphi^0_{t_1}$ and $d\varphi^1_{t_1}$ reversed, as well as slightly different constants (compare (119) - (122) below).

Instead of considering the running minimum process m, we have to monitor from time t_1 on the (local) *running maximum* process M which is defined by

$$M_t = \max_{t_1 \leqslant u \leqslant t} S_u, \qquad t \geqslant t_1.$$

We then define, for $t \geqslant t_1$, similarly as in (104),

$$\tilde{S}_t = \frac{M_t}{\bar{s}}\, g\left(\frac{\bar{s}S_t}{M_t}\right), \tag{106}$$

so that $\tilde{S}_t = (1-\lambda)S_t$ whenever $S_t = M_t$, in which case we sell stock in infinitesimal portions of order $dt^{1/2}$. When $S_t < M_t$ we have $\tilde{S}_t > (1-\lambda)S_t$ in (106) and we do not do any trading. We continue to act according to these rules until the next "large" *negative* excursion happens where we get $S_t = \frac{M_t}{\bar{s}}$ so that $\tilde{S}_t = S_t$ in (106). When this happens we again switch to the regime of buying stock, monitoring (locally) the running minimum process m_t etc etc.

We repeatedly used the word *"locally"* when speaking about the running minimum $(m_t)_{t\geqslant 0}$ (resp. running maximum $(M_t)_{t\geqslant 0}$) of S_t. Let us make precise what we have in mind, thus also starting to translate the above heuristics (e.g., arguing with "immediate" excursions) into proper mathematics. At time $t = 0$, we start by defining $\tilde{S}_0 := S_0$ which corresponds to the fact that we assume that at time $t = 0$ the agent buys stock (which holds true for $\mu > 0$ and λ sufficient small).

Now define sequences of stopping times $(\varrho_n)_{n=0}^{\infty}, (\sigma_n)_{n=1}^{\infty}$ and processes $(m_t)_{t\geqslant 0}$ and $(M_t)_{t\geqslant 0}$ as follows: let $\varrho_0 = 0$ and m the running minimum process of S, i.e.

$$m_t = \inf_{\varrho_0 \leqslant u \leqslant t} S_u, \qquad 0 \leqslant t \leqslant \sigma_1, \tag{107}$$

where the stopping time σ_1 is defined as

$$\sigma_1 = \inf\{t \geqslant \varrho_0 : \tfrac{S_t}{m_t} \geqslant \bar{s}\}.$$

Next define M as the running maximum process of S after time σ_1, i.e.

$$M_t = \sup_{\sigma_1 \leqslant u \leqslant t} S_u, \qquad \sigma_1 \leqslant t \leqslant \varrho_1, \tag{108}$$

where the stopping time ϱ_1 is defined as

$$\varrho_1 = \inf\{t \geqslant \sigma_1 : \tfrac{S_t}{M_t} \leqslant \tfrac{1}{\bar{s}}\}.$$

For $t \geqslant \varrho_1$, we again define

$$m_t = \inf_{\varrho_1 \leqslant u \leqslant t} S_u, \qquad \varrho_1 \leqslant t \leqslant \sigma_2, \tag{109}$$

where

$$\sigma_2 = \inf\{t \geqslant \varrho_1 : \tfrac{S_t}{m_t} \geqslant \bar{s}\},$$

and, for $t \geqslant \sigma_2$, we define

$$M_t = \sup_{\sigma_2 \leqslant u \leqslant t} S_u, \qquad \sigma_2 \leqslant t \leqslant \varrho_2,$$

where

$$\varrho_2 = \inf\{t \geqslant \sigma_2 : \tfrac{S_t}{M_t} \leqslant \tfrac{1}{\bar{s}}\}.$$

Continuing in an obvious way we obtain a.s. finite stopping times $(\varrho_n)_{n=0}^{\infty}$ and $(\sigma_n)_{n=1}^{\infty}$, increasing a.s. to infinity, such that m (resp. M) are the relative running minima (resp. maxima) of S defined on the stochastic intervals $(\llbracket \varrho_{n-1}, \sigma_n \rrbracket)_{n=1}^{\infty}$ (resp. $(\llbracket \sigma_n, \varrho_n \rrbracket)_{n=1}^{\infty}$). Note that

$$\bar{s} m_{\varrho_n} = M_{\varrho_n} = \bar{s} S_{\varrho_n}, \qquad \text{for } n \in \mathbb{N},$$

and

$$\bar{s} m_{\sigma_n} = M_{\sigma_n} = S_{\sigma_n}, \qquad \text{for } n \in \mathbb{N}.$$

We may therefore continuously extend the processes m and M to $\mathbb{R}_+$ by letting

$$M_t := \bar{s} m_t, \qquad \text{for } t \in \bigcup_{n=0}^{\infty} \llbracket \varrho_n, \sigma_{n+1} \rrbracket, \tag{110}$$

$$m_t := \tfrac{M_t}{\bar{s}}, \qquad \text{for } t \in \bigcup_{n=1}^{\infty} \llbracket \sigma_n, \varrho_n \rrbracket. \tag{111}$$

For $t \geqslant 0$, we then have $\bar{s} m_t = M_t$ as well as $m_t \leqslant S_t \leqslant M_t$, and hence

$$m_t \leqslant S_t \leqslant \bar{s} m_t, \qquad \text{for } t \geqslant 0.$$

By construction, the processes m and M are of finite variation and only decrease (resp. increase) on the predictable set $\{m_t = S_t\}$ (resp. $\{M_t = S_t\} = \{m_t = S_t / \bar{s}\}$).

We thus have that the process

$$X_t = \frac{S_t}{m_t} = \frac{\bar{s} S_t}{M_t} \tag{112}$$

takes values in $[1, \bar{s}]$, is reflected at the boundaries and satisfies

$$dX_t = X_t(\mu\, dt + \sigma\, dW_t), \tag{113}$$

when $X_t \in \;]1, \bar{s}[$.

In other words, $([m_t, M_t])_{t \geqslant 0}$ is an interval-valued process such that $\frac{M_t}{m_t} \equiv \bar{s}$, and such that S_t always lies in $[m_t, M_t]$. The interval $([m_t, M_t])_{t \geqslant 0}$ only changes location when S_t touches m_t or M_t, in which case m_t is driven down (resp. M_t is driven up) whenever S_t hits m_t (resp. M_t).

The full SDE satisfied by the process X therefore is

$$dX_t = X_t(\mu\, dt + \sigma\, dW_t) - \frac{dm_t}{m_t} \left(\mathbb{1}_{\{X_t = 1\}} + \bar{s} \mathbb{1}_{\{X_t = \bar{s}\}} \right). \tag{114}$$

3.5 Formulation of the Theorem

Finally, it is time to formulate a mathematically precise theorem.

Theorem 3.6. *Fix $\theta = \frac{\mu}{\sigma^2} \in \,]0,1[$ and $S_0 = 1$ in the Black-Scholes model (69). Let $c(\lambda), \bar{s}(\lambda)$, and $g(\cdot) = g_{c(\lambda)}(\cdot)$ be as in Proposition 3.5 where we suppose that the transaction costs $\lambda > 0$ are sufficiently small.*

Define the continuous process $\tilde{S} = (\tilde{S}_t)_{t \geqslant 0}$ by

$$\tilde{S}_t = m_t \; g\left(\frac{S_t}{m_t}\right), \qquad t \geqslant 0, \tag{115}$$

where the process $(m_t)_{t \geqslant 0}$ is defined in (107), (109) and (111).

Then $\tilde{S}$ is an Itô process, starting at $\tilde{S}_0 = 1$, and satisfying the stochastic differential equation

$$d\tilde{S}_t = g'\left(\frac{S_t}{m_t}\right) \; dS_t + \frac{1}{2m_t} \; g''\left(\frac{S_t}{m_t}\right) \; d\langle S\rangle_t. \tag{116}$$

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

Proof: We may apply Itô's formula to (115). Using (112), (114) and keeping in mind that $(m_t)_{t \geqslant 0}$ is of finite variation, we obtain

$$\begin{aligned}
d\tilde{S}_t &= d\left(m_t g(X_t)\right) \\
&= m_t \; d\left(g(X_t)\right) + g(X_t) \; dm_t \\
&= m_t \left(g'(X_t) \; dX_t + \tfrac{g''(X_t)}{2} \; d\langle X\rangle_t\right) + g(X_t) dm_t \\
&= m_t \Big(g'(X_t) \Big(X_t(\mu \; dt + \sigma \; dW_t) - \tfrac{dm_t}{m_t} \left(\mathbb{1}_{\{X_t = 1\}} + \bar{s}\mathbb{1}_{\{X_t = \bar{s}\}}\right)\Big) \\
&\qquad\qquad + \tfrac{g''(X_t)}{2} \; X_t^2 \sigma^2 \; dt\Big) + g(X_t) \; dm_t \\
&= g'\left(\tfrac{S_t}{m_t}\right) S_t(\mu \; dt + \sigma \; dW_t) + \tfrac{1}{2} g''(\tfrac{S_t}{m_t}) \tfrac{1}{m_t} \; S_t^2 \sigma^2 \; dt \\
&\quad - g'(X_t) \; dm_t \left(\mathbb{1}_{\{X_t = 1\}} + \bar{s}\mathbb{1}_{\{X_t = \bar{s}\}}\right) + g(X_t) \; dm_t \\
&= g'\left(\tfrac{S_t}{m_t}\right) \; dS_t + \frac{g''\left(\frac{S_t}{m_t}\right)}{2m_t} \; d\langle S\rangle_t,
\end{aligned}$$

where in the last line we have used that $dm_t \neq 0$ only on $\{X_t = 1\} \cup \{X_t = \bar{s}\}$ and $g(s) = s \; g'(s)$ for $s = 1$ as well as for $s = \bar{s}$. ∎

Corollary 3.7. *Under the assumptions of Theorem 3.6, fix a horizon $T > 0$ and consider an economic agent with initial endowment $x > 0$ who can trade in a frictionless way in the stock $(\tilde{S}_t)_{0 \leqslant t \leqslant T}$ as defined in (115).*

The unique process $(\hat{\varphi}^0_t, \hat{\varphi}^1_t)_{0\leqslant t\leqslant T}$ *of holdings in bond and stock respectively which optimizes*

$$\mathbb{E}\left[\log\left(x + (\varphi^1 \cdot \tilde{S})_T\right)\right] \to \max! \tag{117}$$

where φ^1 *runs through all predictable,* $\tilde{S}$*-integrable, admissible*[1] *processes and* $\varphi^0_t = x + (\varphi^1 \cdot \tilde{S})_t - \varphi^1_t \tilde{S}_t$*, is given by the following formulae.*

$$(\hat{\varphi}^0_{0-}, \hat{\varphi}^1_{0-}) = (x, 0), \quad (\hat{\varphi}^0_0, \hat{\varphi}^1_0) = \left(\frac{c}{c+1}x, \frac{1}{c+1}x\right) \tag{118}$$

and

$$\hat{\varphi}^0_t = \hat{\varphi}^0_{\varrho_{k-1}} \left(\frac{m_t}{m_{\varrho_{k-1}}}\right)^{\frac{1}{c+1}} \text{ on } \bigcup_{k=1}^{\infty} [\![\varrho_{k-1}, \sigma_k]\!], \tag{119}$$

$$\hat{\varphi}^0_t = \hat{\varphi}^0_{\sigma_k} \left(\frac{m_t}{m_{\sigma_k}}\right)^{\frac{(1-\lambda)\bar{s}}{c+(1-\lambda)\bar{s}}} \text{ on } \bigcup_{k=1}^{\infty} [\![\sigma_k, \varrho_k]\!], \tag{120}$$

as well as

$$\hat{\varphi}^1_t = \hat{\varphi}^1_{\varrho_{k-1}} \left(\frac{m_t}{m_{\varrho_{k-1}}}\right)^{-\frac{c}{c+1}} \text{ on } \bigcup_{k=1}^{\infty} [\![\varrho_{k-1}, \sigma_k]\!], \tag{121}$$

$$\hat{\varphi}^1_t = \hat{\varphi}^1_{\sigma_k} \left(\frac{m_t}{m_{\sigma_k}}\right)^{-\frac{c}{c+(1-\lambda)\bar{s}}} \text{ on } \bigcup_{k=1}^{\infty} [\![\sigma_k, \varrho_k]\!]. \tag{122}$$

The corresponding fraction of wealth invested into stock is given by

$$\tilde{\pi}_t = \frac{\hat{\varphi}^1_t \tilde{S}_t}{\hat{\varphi}^0_t + \hat{\varphi}^1_t \tilde{S}_t} = \frac{1}{1 + c/g(\frac{S_t}{m_t})}. \tag{123}$$

Proof: By (115), $\tilde{S}$ is an Itô process with locally bounded coefficients. We may write (116) as

$$\begin{aligned}\frac{d\tilde{S}_t}{\tilde{S}_t} &= g'\left(\frac{S_t}{m_t}\right)\frac{dS_t}{m_t g(\frac{S_t}{m_t})} + \frac{1}{2m_t^2} g''\left(\frac{S_t}{m_t}\right)\frac{d\langle S\rangle_t}{g(\frac{S_t}{m_t})} \\ &= \underbrace{\frac{S_t^2\sigma^2 g'(\frac{S_t}{m_t})^2}{m_t^2\left(c + g(\frac{S_t}{m_t})\right)g(\frac{S_t}{m_t})}}_{=:\tilde{\mu}_t} dt + \underbrace{\frac{S_t\sigma g'(\frac{S_t}{m_t})}{m_t g(\frac{S_t}{m_t})}}_{=:\tilde{\sigma}_t} dW_t\end{aligned}$$

[1] Admissibility of φ^1 is defined by requiring that the stochastic integral $\varphi^1 \cdot \tilde{S}$ remains uniformly bounded from below.

It follows from the ODE (85) that the mean variance ratio process $\frac{\tilde{\mu}_t}{\tilde{\sigma}_t^2}$ is a bounded process given by

$$\frac{\tilde{\mu}_t}{\tilde{\sigma}_t^2} = \frac{1}{1 + c/g(\frac{S_t}{m_t})}. \tag{124}$$

On the other hand, the adapted process $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t\geqslant 0}$ defined in (119) - (122) is predictable. By definition,

$$\hat{\varphi}_t^0 = cm_t\hat{\varphi}_t^1, \quad t \geqslant 0. \tag{125}$$

For any $k \in \mathbb{N}$, Itô's formula, equation (125), and the fact that $dm_t \neq 0$ only on $\{S_t = m_t\}$ yield

$$d\hat{\varphi}_t^0 + \tilde{S}_t d\hat{\varphi}_t^1 = \left[\left(\frac{m_t}{m_{\varrho_{k-1}}}\right)^{-c/(c+1)} \frac{1}{c+1}\left(\frac{\hat{\varphi}^0_{\varrho_{k-1}}}{m_{\varrho_{k-1}}} - c\hat{\varphi}^1_{\varrho_{k-1}}\right)\right] dm_t = 0,$$

on $[\![\rho_{k-1}, \sigma_k]\!]$ and likewise on $[\![\sigma_k, \rho_k]\!]$ where we use the fact that $dm_t \neq 0$ only on $\{S_t = \bar{s}m_t\}$. Therefore $(\hat{\varphi}^0, \hat{\varphi}^1)$ is self-financing. Again by (125), the fraction

$$\frac{\hat{\varphi}_t^1\tilde{S}_t}{\hat{\varphi}_t^0 + \hat{\varphi}_t^1\tilde{S}_t} = \frac{1}{1 + c/g(\frac{S_t}{m_t})}$$

of wealth invested into stocks, when following $(\hat{\varphi}^0, \hat{\varphi}^1)$, coincides with the Merton proportion computed in (124). Hence $(\hat{\varphi}^0, \hat{\varphi}^1)$ is log-optimal and we are done. ■

In order to discuss the economic message of Corollary 3.7, it is instructive to – formally – pass to the limiting case $\lambda = 0$. In this case we have $\tilde{S}_t = S_t = m_t = M_t$, as well as $c = \frac{1-\theta}{\theta}$ and $\bar{s} = 1$, so that the exponents in (119) - (122) equal

$$\frac{1}{c+1} = \theta, \quad -\frac{c}{c+1} = \theta - 1.$$

We thus find after properly passing to the limits in (119) - (122) the well known formulae due to R. Merton [71]

$$\hat{\varphi}_t^0 = (1-\theta)S_t^\theta, \qquad \hat{\varphi}_t^1 = \theta S_t^{\theta-1} \tag{126}$$

and the fraction of wealth $\tilde{\pi}_t$ invested into stock equals

$$\tilde{\pi}_t = \frac{1}{c+1} = \theta. \tag{127}$$

Passing again to the present case $\lambda > 0$, we have $c > \frac{1-\theta}{\theta}$ and $\bar{s} > 1$. We then find for the exponents in (119), (120)

$$\frac{1}{c+1} < \theta < \frac{(1-\lambda)\bar{s}}{c + (1-\lambda)\bar{s}}. \tag{128}$$

In fact, as was kindly pointed out to us by Paolo Guasoni ([39, Remark after Theorem 5.1]) θ is precisely the arithmetic mean of $\frac{1}{1+c}$ and $\frac{(1-\lambda)\bar{s}}{c+(1-\lambda)\bar{s}}$; this fact can be verified by inserting the formulae (87), (88), (90), and (93) into the identity $g(\bar{s}) = (1-\lambda)\bar{s}$ (compare [34]).

The economic message of (119) - (122) is that we now have to distinguish between the intervals $[\![\varrho_{k-1}, \sigma_k]\!]$ and $[\![\sigma_k, \varrho_k]\!]$. The former are those periods of time when $(m_t)_{0\leqslant t\leqslant T}$ is non-increasing; correspondingly during these intervals the agent only buys stock so that $(\varphi_t^0)_{0\leqslant t\leqslant T}$ is decreasing and $(\varphi_t^1)_{0\leqslant t\leqslant T}$ is increasing. Similarly, the intervals $[\![\sigma_k, \varrho_k]\!]$ are those periods during which $(m_t)_{0\leqslant t\leqslant T}$ is non-decreasing so that the agent only sells stock. The dependence (126) of $(\varphi_t^0)_{0\leqslant t\leqslant T}$ and $(\varphi_t^1)_{0\leqslant t\leqslant T}$ on $\tilde{S}_t = S_t = m_t$ via a power of this process now is replaced by the equations (119) - (122) where the exponents are somewhat different from θ and $(1-\theta)$ respectively, and where we have to distinguish whether we are in the buying or in the selling regime.

As regards the fraction of wealth $\tilde{\pi}_t$ invested into the stock $\tilde{S}$, the message of (123) is that this fraction oscillates between $\frac{1}{1+c}$ and $\frac{1}{1+c/((1-\lambda)\bar{s})}$ as $X_t = \frac{S_t}{m_t}$ oscillates between 1 and $\bar{s}$. Looking again at (128) we obtain — thanks to Paolo Guasoni's observation — that the Merton proportion θ lies precisely in the middle of these two quantities. Economically speaking, this means that the no-trade region is perfectly symmetric around θ, provided that we measure it in terms of the fraction $\tilde{\pi}_t$ of wealth invested into stock where we value the stock by the shadow price $\tilde{S} = g(s)$.

The most important message of Corollary 3.7 is that the optimal strategy $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{0\leqslant t\leqslant T}$ only moves when $(m_t)_{t\geqslant 0}$ moves; the buying of stock takes place when $\tilde{S}_t = S_t$ while selling happens only when $\tilde{S}_t = (1-\lambda)S_t$. This property will be crucial when interpreting $\tilde{S}$ as a shadow price process for the bid-ask process $([(1-\lambda)S_t, S_t])_{0\leqslant t\leqslant T}$.

Another important feature of the present situation is time homogeneity. The conclusion of Corollary 3.7 does not depend on the horizon T.

3.6 Formulation of the optimization problem

We now know that Corollary 3.7 is the answer. But we don't know yet precisely, what the question is! To prepare for the precise formulation, let us start with a formal definition of admissible trading strategies in the presence of transaction costs $\lambda > 0$.

Definition 3.8. *Fix a strictly positive stock price process $S = (S_t)_{0 \leqslant t \leqslant T}$ with continuous paths and transaction costs $\lambda > 0$.*

A self-financing trading strategy starting with zero endowment *is a pair of right continuous, adapted finite variation processes $(\varphi^0_t, \varphi^1_t)_{0 \leqslant t \leqslant T}$ such that*
(i) $\varphi^0_{0-} = \varphi^1_{0-} = 0$
(ii) $\varphi^0_t = \varphi^{0,\uparrow}_t - \varphi^{0,\downarrow}_t$ *and* $\varphi^1_t = \varphi^{1,\uparrow}_t - \varphi^{1,\downarrow}_t$, *where* $\varphi^{0,\uparrow}_t, \varphi^{0,\downarrow}_t, \varphi^{1,\uparrow}_t$, *and* $\varphi^{1,\downarrow}_t$ *are the decompositions of φ^0 and φ^1 into the difference of increasing processes, starting at $\varphi^{0,\uparrow}_{0-} = \varphi^{0,\downarrow}_{0-} = \varphi^{1,\uparrow}_{0-} = \varphi^{1,\downarrow}_{0-} = 0$, and satisfying*

$$d\varphi^{0,\uparrow}_t \leqslant (1-\lambda) S_t d\varphi^{1,\downarrow}_t, \quad d\varphi^{0,\downarrow}_t \geqslant S_t d\varphi^{1,\uparrow}_t, \quad 0 \leqslant t \leqslant T. \tag{129}$$

The trading strategy (φ^0, φ^1) is called admissible *if there is $M > 0$ such that*

$$V_t(\varphi^0, \varphi^1) := \varphi^0_t + (\varphi^1_t)^+ (1-\lambda) S_t - (\varphi^1_t)^- S_t \geqslant -M, \tag{130}$$

holds true a.s., for $0 \leqslant t \leqslant T$.

For example, the process $(\hat{\varphi}^0_t - x, \hat{\varphi}^1_t)_{0 \leqslant t \leqslant T}$, where $(\hat{\varphi}^0, \hat{\varphi}^1)$ was defined in Corollary 3.7 is an admissible trading strategy with zero endowment. Indeed, the buying of the stock, i.e. $d\varphi^{1,\uparrow}_t \neq 0$, only takes place when $\tilde{S}_t = S_t$ and the selling, i.e. $d\varphi^{1,\downarrow}_t \neq 0$, happens only when $\tilde{S}_t = (1-\lambda) S_t$. In addition, $(\hat{\varphi}^0_t)_{t \geqslant 0}$ and $(\hat{\varphi}^1_t)_{t \geqslant 0}$ are of finite variation and as $0 < \theta < 1$, we have $\hat{\varphi}^0_t > 0, \hat{\varphi}^1_t > 0$.

Now we define a convenient version of our optimization problem.

Definition 3.9. *Fix $\theta = \frac{\mu}{\sigma^2} \in \left]0,1\right[$ in the Black-Scholes model* (69)*, transaction costs $\lambda > 0$ sufficiently small, as well as an initial endowment $x > 0$ and a horizon T.*

Let $(\tilde{S}_t)_{0 \leqslant t \leqslant T}$ be the process defined in Theorem 3.6. The optimization problem is defined as

$$(P_x) \qquad \mathbb{E}\left[\log(x + \varphi^0_T + \varphi^1_T \tilde{S}_T)\right] \to \max! \tag{131}$$

where (φ^0, φ^1) runs through the admissible trading strategies with transaction costs λ starting with zero endowment $(\varphi^0_{0\_}, \varphi^1_{0\_}) = (0,0)$.

The definition is designed in such a way that the subsequent result holds true.

Theorem 3.10. *Under the hypotheses of Definition 3.9 the unique optimizer in* (131) *is $(\hat{\varphi}^0 - x, \hat{\varphi}^1)$, where $(\hat{\varphi}^0, \hat{\varphi}^1)$ are given by Corollary 3.7.*

Proof: The process $(\hat{\varphi}^0, \hat{\varphi}^1)$ is the unique optimizer to the optimization problem (117) when we optimize over the larger class of admissible trading strategies in the frictionless market $\tilde{S}$.

As $(\hat{\varphi}^0, \hat{\varphi}^1)$ also is an admissible trading strategy in the sense of Definition 3.8 the assertion of the theorem follows *a fortiori*. ■

Let us have a critical look at the precise features of Definition 3.9. After all, we are slightly cheating: we use the process $\tilde{S}$, which is *part of the solution*, for the *formulation of the problem.* Why do we do this trick? We just have seen that this way of defining the optimization problem allows for the validity of the elegant Theorem 3.10. We also remark that Theorem 3.10 exhibits the same time homogeneity, i.e. non-dependence on the horizon T, as Theorem 3.1 and Corollary 3.7.

But the honest formulation of problem (131) would be

$$(P'_x) \qquad \mathbb{E}\left[\log(x + \varphi^0_T + (\varphi^1_T)^+(1-\lambda)S_T - (\varphi^1_T)^- S_T)\right] \to \max! \tag{132}$$

The economic interpretation of (P'_x) is that at time T the liquidation of the position φ^1_T in stock has to be done at the ask price S_T or the bid price $(1-\lambda)S_T$, depending on the sign of φ^1_T. On the other hand the problem (P_x) in (131) allows for liquidation at the shadow price $\tilde{S}_T$, which is a random variable taking values in $[(1-\lambda)S_T, S_T]$.

The problem (P'_x) does not allow for a mathematically nice treatment as it lacks time homogeneity (see [34] for a more detailed discussion pertaining to the economic aspects). But (P_x) is a good proxy for (P'_x): the difference between $\tilde{S}_T$ as opposed to $(1-\lambda)S_T$ and S_T is of order λ and only pertains to *one instance* of trading, namely at time T. On the other hand we have seen in Proposition 3.5 (compare also Proposition 3.11 below) that the leading terms of the effects of transaction costs on the *dynamic trading* activities during the interval $[0, T[$ are of order $\lambda^{1/3}$. Hence, for fixed horizon T, the latter effect becomes dominant as $\lambda \to 0$.

The situation becomes even better if we consider the limiting case $T \to \infty$. After proper normalization (see, e.g., (135) below) the difference between (P_x) and (P'_x) completely disappears in the limit $T \to \infty$. For example, in (137) below we find the exact dependence on $\lambda > 0$ (involving *all* the powers of $\lambda^{1/3}$) independently of whether we consider the problem (P_x) or (P'_x). For all these reasons we believe that (P_x) is the "good" definition of the problem.

3.7 The Case $\theta \geqslant 1$

The preceding results pertain to the case $0 < \theta < 1$, where we have seen that the optimal holdings $(\hat{\varphi}^0_t, \hat{\varphi}^1_t)$ in bond as well as in stock are strictly positive, for all $t \geqslant 0$.

The case $\theta = 1$ is degenerate. As is well known and immediately deduced from Theorem 3.1, in the absence of transaction costs the optimal strategy consists in fully investing the initial endowment x into stock at time zero, so that $\hat{\varphi}_t^0 \equiv 0$ and $\hat{\varphi}_t^1 \equiv x$, if S_0 is normalized to 1. In the presence of transaction costs $\lambda > 0$ it is rather obvious, from an economic point of view, that this strategy still is optimal. In fact, if we define the shadow price process $\tilde{S}$ simply by $\tilde{S}_t = S_t$, then the above strategy $(\hat{\varphi}_t^0, \hat{\varphi}_t^1) = (0, x)$, for $0 \leqslant t \leqslant T$ also is the solution to the problem (P_x) in (131) in a formal way.

More challenging is the case $\theta > 1$. In this regime the well-known frictionless optimal strategy involves a *short position* in bond, i.e. $\varphi_t^0 < 0$, and using this leverage to finance a long position φ_t^1 in stock, so that $\varphi_t^1 S_t$ exceeds the current wealth of the agent.

This phenomenon also carries over to the situation under (sufficiently small) transaction costs $\lambda > 0$. In this situation the agent *buys* stock when stock prices are rising and *sells* stock when stock prices are falling, i.e., she has the opposite behavior of the case $0 < \theta < 1$.

Mathematically speaking, this results in the fact that we again look at the function g as defined in (87), satisfying the ODE (85), but now the domain of definition of g is given by an interval $[\bar{s}, 1]$, where $\bar{s} < 1$.

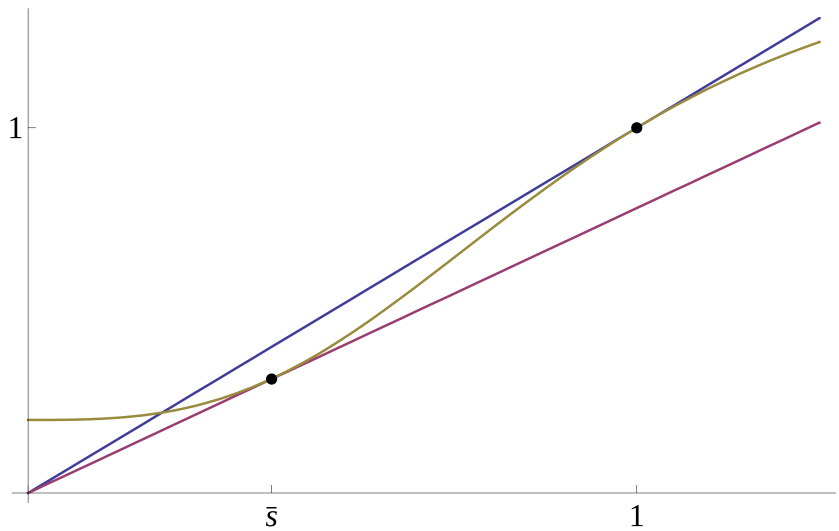

Figure 6: Smooth pasting conditions for the function g, for $\theta > 1$.

The boundary conditions still are given by (86) and (89), and the formula for $c = c(\lambda)$ and $\bar{s} = \bar{s}(\lambda)$ still are given by (97) and (98) (applying the convention $(-x)^{1/3} = -(x^{1/3})$, for $x > 0$; see [34, Prop. 6.1] for details).

Hence, also in the case $\theta > 1$ we find an analogous situation as for $0 < \theta < 1$. The story simply has to be told the other way round: we start again with the normalizing assumption $S_0 = 1$, as well as the definition $\tilde{S}_0 = g(S_0) = 1$, which corresponds to assuming that the agent buys stock at time $t = 0$, just as above.

Now suppose (heuristically) that the stock starts a *negative* excursion at time $t = 0$, i.e. $S_t < 1$, for $t > 0$ small enough. We then define $\tilde{S}$ by

$$\tilde{S}_t = g(S_t), \qquad t \geqslant 0,$$

up to time $t_0 > 0$ when S_t hits again 1, or when S_t hits for the first time $\bar{s}$ (which now is less than 1).

Passing to the general (and generic) case, i.e. dropping the assumption about the negative excursion starting at $t = 0$, we define the running maximum process $(M_t)_{t \geqslant 0}$ locally by

$$M_t = \sup_{0 \leqslant u \leqslant t} S_u, \qquad 0 \leqslant t \leqslant \varrho_1$$

where ϱ_1 is the first time when $S_t/M_t \leqslant \bar{s}$. We define

$$\tilde{S}_t = M_t g\left(\frac{S_t}{M_t}\right), \qquad \text{for } 0 \leqslant t \leqslant \varrho_1.$$

During the stochastic interval $[\![0, \varrho_1]\!]$ the agent *buys* stock whenever $(M_t)_{0 \leqslant t \leqslant \varrho_1}$ moves up, following a similar logic as in (119) - (122) above.

After time ϱ_1 the agent monitors locally the running minimum process $(m_t)_{t \geqslant \varrho_1}$

$$m_t = \min_{\varrho_1 \leqslant u \leqslant t} S_u, \qquad \varrho_1 \leqslant t \leqslant \sigma_1$$

where σ_1 is the first time when $\frac{S_t}{m_t} \geqslant \frac{1}{\bar{s}}$. We define $\tilde{S}_t := \frac{m_t}{\bar{s}} g(\frac{\bar{s} S_t}{m_t})$ for $\varrho_1 \leqslant t \leqslant \sigma_1$. During the stochastic interval $[\![\varrho_1, \sigma_1]\!]$, the agent *sells* stock when m_t moves down.

The reasoning is perfectly analogous to section 3.4 above. We refer to [34] for details and only mention that, for $\theta > 1$, the parameter c in Proposition 3.4 now has to vary in $]\frac{1-\theta}{\theta}, 0[$.

There is still one slightly delicate issue in the case $\theta > 1$ which we have not yet discussed: the *admissibility* of the optimal strategies $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t \geqslant 0}$ which, also in the case $\theta > 1$, are given by formulas (118) - (122). Now the holdings $(\hat{\varphi}_t^1)_{t \geqslant 0}$ in bond are negative so that we have to check more carefully whether the agent is solvent at all times $t \geqslant 0$. As $\hat{\varphi}_t^1 \geqslant 0$, the natural condition is

$$\hat{\varphi}_t^0 + \hat{\varphi}_t^1 S_t (1 - \lambda) \geqslant 0, \qquad t \geqslant 0. \tag{133}$$

We know that

$$\hat{\varphi}_t^0 + \hat{\varphi}_t^1 \tilde{S}_t \geqslant 0, \tag{134}$$

a.s., for each $t \geqslant 0$. Indeed $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t \geqslant 0}$ is the log-optimal portfolio for the frictionless market $(\tilde{S}_t)_{t \geqslant 0}$; it is well-known form the frictionless theory (Theorem 3.1) and rather obvious that (134) has to hold true.

To show that even (133) is satisfied, fix $t_0 \geqslant 0$ and $(\hat{\varphi}_{t_0}^0, \hat{\varphi}_{t_0}^1, \tilde{S}_{t_0})$ such that $\tilde{S}_{t_0} \in \,](1-\lambda)S_{t_0}, S_{t_0}[$. Conditionally on $(\hat{\varphi}_{t_0}^0, \hat{\varphi}_{t_0}^1, \tilde{S}_{t_0})$ define the stopping times ϱ and σ.

$$\begin{aligned} \varrho &= \inf\{t > t_0 : \tilde{S}_t = S_t\}, \\ \sigma &= \inf\{t > t_0 : \tilde{S}_t = (1-\lambda)S_t\}. \end{aligned}$$

Clearly we have, conditionally on $(\hat{\varphi}_{t_0}^0, \hat{\varphi}_{t_0}^1, \tilde{S}_{t_0})$, that $\mathbb{P}[\sigma < \varrho] > 0$. As $(\hat{\varphi}_t^0, \hat{\varphi}_t^1)_{t_0 \leqslant t \leqslant \sigma \wedge \varrho}$ remains constant and using $S_\sigma < S_{t_0}$ on $\{\sigma < \varrho\}$ we deduce from

$$\hat{\varphi}_\sigma^0 + \hat{\varphi}_\sigma^1 \tilde{S}_\sigma \geqslant 0, \qquad \text{on } \{\sigma < \varrho\}$$

that

$$\hat{\varphi}_\sigma^0 + \hat{\varphi}_\sigma^1 (1-\lambda) S_\sigma \geqslant 0 \qquad \text{on } \{\sigma < \varrho\}$$

so that

$$\hat{\varphi}_{t_0}^0 + \hat{\varphi}_{t_0}^1 (1-\lambda) S_{t_0} \geqslant \hat{\varphi}_\sigma^0 + \hat{\varphi}_\sigma^1 (1-\lambda) S_\sigma \geqslant 0.$$

This proves (133).

3.8 The Optimal Growth Rate

We now want to compute the optimal growth rate

$$\delta := \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\log(1 + \hat{\varphi}_T^0 + \tilde{S}_T \hat{\varphi}_T^1)\right] = \limsup_{T\to\infty} \frac{1}{T} \mathbb{E}\left[\int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right], \tag{135}$$

where the initial endowment x is normalized by $x = 1$, and $(\hat{\varphi}^0, \hat{\varphi}^1)$ denotes the log-optimal portfolio for the shadow price $\tilde{S}$ from Corollary 3.7. The second equality follows from Theorem 3.1 and Theorem 3.10 (compare [60, Example 6.4]).

By the construction in (112) the process $X = S/m$ is a geometric Brownian motion with drift which is reflected on the boundaries of the interval $[1, \bar{s}]$ (resp. on $[\bar{s}, 1]$ for the case $\theta > 1$). Therefore, an ergodic theorem for positively recurrent one-dimensional diffusions (cf. e.g. [4, Sections II.36 and II.37]) and elementary integration yield the following result.

Proposition 3.11. *Suppose the conditions of Theorem 3.6 hold true. Then the process $X = S/m$ has the stationary distribution*

$$\nu(ds) = \begin{cases} \dfrac{2\theta - 1}{\bar{s}^{2\theta-1} - 1} s^{2\theta-2} \mathbf{1}_{[1,\bar{s}]}(s) ds & \text{for } \theta \in (0,1)\backslash\{\frac{1}{2}\}, \\ \dfrac{1}{\log(\bar{s})} s^{-1} \mathbf{1}_{[1,\bar{s}]}(s) ds & \text{for } \theta = \frac{1}{2}, \\ \dfrac{2\theta - 1}{1 - \bar{s}^{2\theta-1}} s^{2\theta-2} \mathbf{1}_{[\bar{s},1]}(s) ds & \text{for } \theta \in (1,\infty). \end{cases}$$

Moreover, the optimal growth rate for the frictionless market with price process $\tilde{S}$ as well as for the market with bid-ask process $[(1-\lambda)S, S]$ is given by

$$\begin{aligned} \delta &= \left| \int_1^{\bar{s}} \frac{\tilde{\mu}^2(s)}{2\tilde{\sigma}^2(s)} \nu(ds) \right| \\ &= \begin{cases} \dfrac{(2\theta-1)\sigma^2 \bar{s}}{2(1+c)(\bar{s} + (-2 - c + 2\theta(1+c))\bar{s}^{2\theta})} & \text{for } \theta \in (0,\infty)\backslash\{\frac{1}{2}, 1\}, \\ \dfrac{\sigma^2}{2(1+c)(1+c-\log \bar{s})} & \text{for } \theta = \frac{1}{2}, \end{cases} \end{aligned} \tag{136}$$

where c and $\bar{s}$ denote the constants from Proposition 3.5.

As $\lambda \to 0$, the optimal growth rate has the asymptotics

$$\delta = \frac{\mu^2}{2\sigma^2} - \left(\frac{3\sigma^3}{\sqrt{128}} \theta^2 (1-\theta)^2 \right)^{2/3} \lambda^{2/3} + O(\lambda^{4/3}). \tag{137}$$

Proof: The calculation of the invariant distribution ν of the process X is an elementary exercise. The remaining calculations are tedious, but elementary too (see [34, Proposition 5.4 and 6.3]). ■

3.9 Primal versus Dual Approach

In the preceding arguments we have developed the solution to the problem of finding the growth-optimal portfolio under transaction costs by using the "dual" approach, which also sometimes is called the "martingale method" (compare the pioneering paper [12] by Cvitanic and Karatzas). Starting from the Black-Scholes model (69), we have considered the "shadow price process" $\tilde{S} = (\tilde{S}_t)_{t \geqslant 0}$ which in the notation of (36) corresponds to

$$\tilde{S}_t = \frac{\hat{Z}_t^1}{\hat{Z}_t^0}. \tag{138}$$

In the present context the density process $(\hat{Z}^0_t)_{t\geqslant 0}$ is given by Girsanov's formula

$$\hat{Z}^0_t = \exp\left(-\int_0^t \frac{\tilde{\mu}_s}{\tilde{\sigma}_s} dW_s - \int_0^t \frac{\tilde{\mu}_s^2}{2\tilde{\sigma}_s^2} ds \right). \tag{139}$$

It is the unique $\mathbb{P}$-martingale with respect to the filtration generated by W and starting at $\hat{Z}^0_0 = 1$, such that the process $\hat{Z}^1_t := \hat{Z}^0_t \tilde{S}_t$ is a $\mathbb{P}$-martingale too. As we have seen in Section 2, this solution of the dual problem can be translated into the solution of the primal problem via the first order conditions (37).

It is worthwhile to spell out explicitly the formulation of the dual problem corresponding to (36). The conjugate function $V(y)$ associated to $U(x) = \log(x)$ by (34) is

$$V(y) = -\log(y) - 1, \qquad\qquad y > 0.$$

Under the assumptions of Corollary 3.7 we define for fixed $T > 0$, in analogy to (36) and using (139),

$$\begin{aligned} v(y) &= \mathbb{E}[V(y\hat{Z}^0_T)] \\ &= -\log(y) - 1 + \mathbb{E}\left[\int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right] \\ &= V(y) + \mathbb{E}\left[\int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right]. \end{aligned}$$

Hence we find as in Theorem 2.3 that $v(y)$ is the conjugate function to the indirect utility function associated to the shadow price process $\tilde{S}$

$$\begin{aligned} u(x) &= \mathbb{E}[U(x\hat{V}_T)] \\ &= \log(x) + \mathbb{E}\left[U\left(\exp\left(\left(\int_0^T \frac{\tilde{\mu}_t}{\tilde{\sigma}_t} dW_t + \int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right)\right)\right]\right. \\ &= U(x) + \mathbb{E}\left[\int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right], \end{aligned}$$

where

$$\hat{V}_T = \left(\exp(\int_0^T \frac{\tilde{\mu}_t}{\tilde{\sigma}_t} dW_t + \int_0^T \frac{\tilde{\mu}_t^2}{2\tilde{\sigma}_t^2} dt\right)$$

denotes the optimal terminal wealth for the frictionless market $\tilde{S}$.

The above considerations pertain to the frictionless complete market $\tilde{S}$; they carry over verbatim to the bid as process $[(1-\lambda)S, S]$ if we use definition (131) for the formulation of the portfolio optimization problem.

Another approach to finding the growth optimal portfolio is to directly attack the primal problem which leads to a Hamilton-Jacobi-Bellman equation for the *value function* associated to the primal problem; in economic terminology this value function (see (141) below) is called the"indirect utility function".

This strain of literature has a longer history than the "dual approach" [12]. In [85] Taksar, Klass and Assaf give a solution to the present problem of finding the growth optimal portfolio, and in [28] Dumas and Luciano solve the same problem for power utility $U(x) = \frac{x^\gamma}{\gamma}, 0 < \gamma < 1$, rather than for $U(x) = \log(x)$. Let us also mention the work of Davis and Norman [23] and Shreve and Soner [83] on optimal consumption which proceeds by the primal method too. We refer to [45] for an account on the ample literature persuing this "primal" method.

We shall present here the approach of [85] and [28]. Our aim is to relate the "primal" and the "dual" approach, thus gaining additional insight into the problem. While in the preceding subsections the mathematics were finally done in a rigorous way, we now content ourselves to more informal and heuristic considerations. We can afford to do so as we have established things rigorously already above.

Fixing the level $\lambda > 0$ of (sufficiently small) transactions costs, the horizon T, and an initial endowment $(\varphi^0, \varphi^1) \in \mathbb{R}^2_+$ in bond[2] and stock, we define

$$u(\varphi^0, \varphi^1, s, T) = \sup\{\mathbb{E}[\log(\varphi^0_T + \varphi^1_T S_T)|S_0 = s]\} \tag{140}$$

where $(\varphi^0_T, \varphi^1_T)$ runs through all pairs of positive $\mathcal{F}_T$-measurable random variables (modeling the holdings in units of bond and stock at time T) which can be obtained by admissible trading (and paying transaction costs λ) as in (129), starting from initial positions $(\varphi^0_{0-}, \varphi^1_{0-}) = (\varphi^0, \varphi^1)$.

The term $(\varphi^0_T + \varphi^1_T S_T)$ in (140) above corresponds to the modeling assumption that the position φ^1_T in stock can be liquidated at time T at price S_T. One might also define (140) by using $(\varphi^0_T + \varphi^1_T(1-\lambda)S_T)$. As observed at the end of sub-section 3.6, this difference will play no role when we eventually pass to the (properly scaled) limit $T \to \infty$, hence we may as well use (140) as is done in [28].

Turning back to a fixed horizon $T > 0$, define, for $0 \leqslant t \leqslant T$, the value function

$$u(\varphi^0, \varphi^1, s, t, T) = \sup\{\mathbb{E}[\log(\varphi^0_T + \varphi^1_T S_T)|S_t = s]\}, \tag{141}$$

[2] in [85] and [28] no short-selling is allowed so that $\varphi^0 \geqslant 0, \varphi^1 \geqslant 0$. Hence we assume, as in these papers, that $\theta = \frac{\mu}{\sigma^2} \in \,]0, 1[$.

where now $(\varphi_T^0, \varphi_T^1)$ range in the random variables which can be obtained, similarly as above, by admissible trading during the period $[t, T]$, and starting at time t_- with holdings $(\varphi_{t_-}^0, \varphi_{t_-}^1) = (\varphi^0, \varphi^1)$.

The idea is to pass, for fixed $t > 0$, to the limit $T \to \infty$ in (141) in order to obtain an indirect utility function $u(\varphi^0, \varphi^1, s, t)$ not depending on the horizon T. But, of course, by blindly passing to this limit we shall typically find $u(\varphi^0, \varphi^1, s, t) \equiv \infty$ which yields no information.

The authors of [85] and [28] therefore *assume* that there is a constant $\delta > 0$ such that, by discounting the value of the portfolio $\varphi_T^0 + \varphi_T^1 S_T$ with the factor $e^{\delta T}$, we *get a finite limit* below.

$$\begin{aligned} u(\varphi^0, \varphi^1, s, t) :&= \lim_{T\to\infty} \sup\{\mathbb{E}[\log(e^{-\delta T}(\varphi_T^0 + \varphi_T^1 S_T))|S_t = s]\} \\ &= \lim_{T\to\infty} \sup\{\mathbb{E}[\log(\varphi_T^0 + \varphi_T^1 S_T)|S_t = s]\} - \delta T. \end{aligned} \tag{142}$$

According to our calculations we already *know* that the above $\delta > 0$ must be the optimal growth rate which we have found in (136). But in the primal approach of [85] and [28], the number $\delta > 0$ is a free parameter which eventually has to be determined by analyzing the boundary conditions of the differential equations related to the indirect utility function $u(\varphi^0, \varphi^1, s, t)$.

To analyze the indirect utility function u, we start by making some simplifications. From definition (142) we deduce that

$$u(\varphi^0, \varphi^1, s) := u(\varphi^0, \varphi^1, s, 0) = u(\varphi^0, \varphi^1, s, t) + \delta t, \qquad \text{for } t \geqslant 0 \tag{143}$$

where the left hand side does not depend on t anymore. We also use the scaling property of the logarithm

$$u(\mu\varphi^0, \mu\varphi^1, s) = u(\varphi^0, \varphi^1, s) + \log(\mu),$$

to reduce to the case where we may normalize φ^0 to be one. To eventually reduce the two remaining variables φ^1 and s to simply one dimension, make the economically obvious observation that the variables φ^1 and s only enter into the function u via the product $\varphi^1 s$. Introducing the new variable $y = \frac{\varphi^1 s}{\varphi^0}$, which describes the ratio of the value of the stock investment to the bond investment, we therefore may write u in (142) as

$$\begin{aligned} u(\varphi^0, \varphi^1, s, t) &= \log(\varphi^0) + h\left(\frac{\varphi^1 s}{\varphi^0}\right) - \delta t \\ &= \log(\varphi^0) + h(y) - \delta t \end{aligned} \tag{144}$$

for some function $h : \mathbb{R}_+ \to \mathbb{R}$ to be determined.

Let us find the Hamilton-Jacobi-Bellman equation satisfied by u. According to the basic principle of stochastic optimization [73], we must have that, for any self-financing $\mathbb{R}^2_+$-valued trading strategy $(\varphi^0_t, \varphi^1_t)_{t\geqslant 0}$, the process $(u(\varphi^0_t, \varphi^1_t, S_t, t))_{t\geqslant 0}$ is a super-martingale, which becomes a true (local) martingale if we plug in the optimal strategy $(\hat{\varphi}^0, \hat{\varphi}^1)$.

First consider the possible control of keeping $(\varphi^0_t, \varphi^1_t) = (\varphi^0, \varphi^1)$ simply constant: this yields via (69), (142) and (144)

$$\begin{aligned} du(\varphi^0_t, \varphi^1_t, S_t, t) &= u_s dS_t + \frac{u_{ss}}{2} d\langle S\rangle_t - \delta dt \\ &= \tfrac{\varphi^1_t}{\varphi^0_t} h'(y_t)(S_t\sigma dW_t + S_t\mu dt) + \tfrac{(\varphi^1_t)^2}{(\varphi^0_t)^2} h''(y_t)\left(\frac{S_t^2\sigma^2}{2} dt\right) - \delta dt, \end{aligned}$$

hence, by taking expectations and using the formal identity $\mathbb{E}[dW_t] = 0$,

$$\mathbb{E}[du(\varphi^0_t, \varphi^1_t, S_t, t)] = \left[S_t\mu \frac{\varphi^1_t}{\varphi^0_t} h'(y_t) + \frac{S_t^2\sigma^2}{2} \frac{(\varphi^1_t)^2}{(\varphi^0_t)^2} h''(y_t) - \delta \right] dt.$$

The term in the bracket has to be non-positive. We know already that, within the no-trade region, it is indeed optimal to keep φ^0_t and φ^1_t constant. Hence, by replacing $y_t = \frac{S_t\varphi^1_t}{\varphi^0_t}$ by the real variable $y > 0$, we expect that the function h will satisfy the ODE

$$h''(y)\frac{y^2\sigma^2}{2} + h'(y)y\mu - \delta = 0, \tag{145}$$

where $y = \frac{\varphi^1 s}{\varphi^0}$ ranges in the no-trade region, which should be a compact interval $[l, r]$ contained in $]0, \infty[$, which we still have to determine.

Equation (145) is an elementary ODE which, by passing to logarithmic coordinates $z = \log(y)$, can be reduced to a linear ODE. In particular, it has a closed form solution. For $\theta = \frac{\mu}{\sigma^2} \in \mathbb{R}_+\backslash\{\frac{1}{2}\}$, the general solution is given by

$$h(y) = \frac{\delta}{\mu - \frac{\sigma^2}{2}} \log(y) + C_1 y^{2\theta-1} + C_2, \tag{146}$$

while for the case $\theta = \frac{\mu}{\sigma^2} = \frac{1}{2}$ we obtain

$$h(y) = \frac{\delta}{\sigma^2} \log(y)^2 + C_1 \log(y) + C_2, \tag{147}$$

where the constants C_1, C_2 still are free.

Plugging (146) into the utility function (144) with $t = 0$ we obtain

$$\begin{aligned} u(\varphi^0, \varphi^1, s) &= \log(\varphi^0) + h(y) &(148)\\ &= \log(\varphi^0) + \frac{\delta}{\mu - \frac{\sigma^2}{2}} \log(y) + C_1 y^{2\theta-1}, &(149) \end{aligned}$$

for $\theta \in \mathbb{R}_+ \backslash \{\frac{1}{2}\}$, and a similar expression is obtained for $\theta = \frac{1}{2}$. We have set $C_2 = 0$ above, as an additive constant does not matter for the indirect utility. The parameters C_1 and δ are still free.

In [85] and [28] the idea is to analyze the above function and to determine the free boundaries l, r, such that $y \in [l, r]$ is the no-trade region, where the indirect utility function is given by (149) above. We therefore have to deal with 4 free parameters and to find boundary conditions, involving again smooth pasting arguments, to determine them.

We refer to [85] and [28] for the further analysis of this delicate free boundary problem. Eventually these authors achieve numerical solutions of the free boundary problem, but do not try to obtain analytical results, e.g., to develop the quantities in fractional Taylor series in $\lambda^{1/3}$ as we have done above.

Our concern of interest is the relation of the primal approach, in particular the ODE (145), with the dual approach, in particular with the shadow price process $\tilde{S}$.

This link is given by the economic idea of the *marginal rate of substitution.* Fix t and suppose that the triple $(\varphi^0, \varphi^1, s)$ is such that $y = \frac{\varphi^1 s}{\varphi^0}$ lies in the no-trade region. The indirect utility then is given by (144). Changing the position φ^0 of holdings in bond from φ^0 to $\varphi^0 + d\varphi^0$, for some small $d\varphi^0$, the indirect utility changes (of first order) by the quantity $u_{\varphi^0} d\varphi^0$, where by differentiating (144) and using (146) we have

$$u_{\varphi^0} = \frac{1}{\varphi^0} - \frac{\delta}{\mu - \frac{\sigma^2}{2}} \frac{1}{y} \frac{y}{\varphi^0} - C_1 \; y^{2\theta - 2} \frac{y}{\varphi^0}.$$

Similarly, changing the position of φ^1 units of stock to $\varphi^1 + d\varphi^1$ units for some small $d\varphi^1$, this change of first order equals $d\varphi^1 \; u_{\varphi^1}$, where

$$u_{\varphi^1} = \frac{\delta}{\mu - \frac{\sigma^2}{2}} \frac{1}{y} \frac{y}{\varphi^1} + C_1 \; y^{2\theta - 2} \frac{y}{\varphi^1}.$$

The natural economic question is the following: what is the price $\tilde{s} = \tilde{s}(\varphi^0, \varphi^1, s)$ for which an economic agent is — of first order — indifferent of buying/selling stock against bond? The obvious answer is that the ratio $\tilde{s} = \frac{d\varphi^0}{d\varphi^1}$ must satisfy the equality $u_{\varphi^0} d\varphi^0 = u_{\varphi^1} d\varphi^1$. In other words, $\tilde{s}$ is given by the "marginal rate of substitution"

$$\tilde{s} = \frac{u_{\varphi^1}(\varphi^0, \varphi^1, s)}{u_{\varphi^0}(\varphi^0, \varphi^1, s)} \tag{150}$$

$$= \frac{\varphi^0}{\varphi^1} \cdot \frac{\frac{\delta}{\mu - \frac{\sigma^2}{2}} + C_1 y^{2\theta - 1}}{(1 - \frac{\delta}{\mu - \frac{\sigma^2}{2}}) - C_1 y^{2\theta - 1}}. \tag{151}$$

This formula for $\tilde{S}$ looks already reminiscent of the function $\tilde{S} = g(s)$ in (87). To make this relation more explicit, recall that we have made the following normalizations in subsection 3.2 above: the variable s ranges in the interval $[1, \bar{s}]$ and the ratio $\frac{\varphi^0}{\varphi^1}$ of holdings in bond and stock equals the parameter c in formula (104), if we have the normalization $m_t = 1$, so that $\tilde{S}_t = g(S_t)$. Hence $y = \frac{\varphi^1 s}{\varphi^0} = \frac{s}{c}$ so that in (151) we get

$$\tilde{s} = G(s) := \frac{\frac{c\delta}{\mu - \frac{\sigma^2}{2}} + C_1 c^{2-2\theta} s^{2\theta - 1}}{(1 - \frac{\delta}{\mu - \frac{\sigma^2}{2}}) - C_1 c^{1-2\theta} s^{2\theta - 1}}, \tag{152}$$

Using the relation

$$\delta = \delta(c) = \frac{(2\theta - 1)\sigma^2 \bar{s}(c)}{2(1 + c)(\bar{s} + (-2 - c + 2\theta(1 + c)\bar{s}^{2\theta})}$$

obtained in (136) above, we conclude that the function $G(\cdot)$ defined in (152) above indeed equals the function g in (87) if we choose the free parameter C_1 properly. As the variable s ranges in the interval $[1, \bar{s}]$, we find that the no trade interval $[l, r]$ for the variable y equals $[\frac{1}{c}, \frac{\bar{s}}{c}]$ and we can use the Taylor expansions in powers of $\lambda^{1/3}$ to explicitly determine the values of these boundaries. We thus can provide explicit formulae for all the quantities involved in the solution of the primal problem where the PDE approach only could give numerical solutions.

We now understand better why we found a *closed form solution* for the ODE (85). As regards the function h solving the ODE (145), there is, of course, the closed form solution (146), as this ODE is linear (after passing to logarithmic coordinates). Therefore the indirect utility u in (144) again is given by an explicit formula. Hence the function $G = g$, which is deduced from the "martingale rate of substitution relation" (150), has to be so too.

3.10 Rogers' qualitative argument

We finish this section by recalling a lovely "back of an envelope calculation" due to Ch. Rogers [76]. It shows that the leading term for the size $\bar{s}(\lambda)-1$ of the no trade region is of the order $\lambda^{\frac{1}{3}}$ (compare (98)) and that the difference of the growth rate $\delta(\lambda)$ obtained in (137) to the frictionless growth rate $\frac{\mu^2}{2\sigma^2}$ is of the order $\lambda^{\frac{2}{3}}$. In fact, these relations were already obtained in the early work of G. Constantinides [11].

The starting point is the rather obvious assumption that, given transaction costs $\lambda > 0$, the log optimal investor will keep the ratio of stock to the total wealth investment in an interval of width w around the Merton proportion $\theta = \frac{\mu}{\sigma^2}$.

Taking the frictionless market as benchmark, what are the (negative) effects of transaction costs λ when choosing the width w? There are two causes. On the one hand side one has to pay transaction costs TRC. From scaling it is rather obvious, at least asymptotically, that these costs are proportional to the size of transaction costs λ and indirectly proportional to the width w, i.e. $TRC \approx c_1\lambda w^{-1}$ for some constant c_1. Indeed, the local time spent at the boundary of the no trade region, where trading takes place, is of the order w^{-1}.

The second negative influence is the cost of misplacement: in comparison to the ideal ratio of the Merton proportion one typically is of the order w away from it. As the utility function attains its optimum at the Merton proportion (and assuming sufficient smoothness), the effect of the misplacement on the performance should be proportional to the square of the misplacement. This is, at least heuristically, rather obvious. Actually, the fact that a function decreases like a square when it is close to its maximum was already observed as early as in 1613 by Johannes Kepler in the context of the volume of wine barrels. Hence the misplacement cost MPC caused by the width w of the no trade region should asymptotically satisfy $MPC \approx c_2 w^2$, for some constant c_2.

The total cost TC of these two causes therefore has an asymptotic bahavior of the form

$$TC = TRC + MPC \approx c_1\lambda w^{-1} + c_2 w^2.$$

We have to minimize this expression as a function of w. Setting the derivative of this function equal to zero gives for the optimal width $\hat{w}$ the asymptotic relation $\hat{w} \approx c\lambda^{1/3}$, where $c = \left(\frac{c_1}{2c_2}\right)^{1/3}$.

As regards the effect of the transaction costs λ on the asymptotic growth rate, we conclude from the above argument that this is the order of the square of the typical misplacement $\hat{\omega}$ which in turn is of the order $\lambda^{1/3}$. Therefore the difference of the frictionless growth rate $\frac{\mu^2}{2\sigma^2}$ to the rate involving transaction costs is of the oder $\lambda^{2/3}$ (compare (137)).