PASSPORT OPTIONS

Freddy Delbaen, Eidgenössische Technische Hochschule, Zürich
Marc Yor, Laboratoire de Probabilités,
Université Pierre et Marie Curie, Paris VI

first version May 8, 1998, this version December 1, 1999

Abstract. We relate the theory of passport options with general principles from martingale theory as well as with the theory of Bessel processes.

1991 *Mathematics Subject Classification.* 90A09, 49L99, 60G44, 60H07, 60J55, 60J65.

Key words and phrases. Bessel processes, financial derivatives, martingale inequalities, Passport option, time transforms.

Part of this research was sponsored by Credit Suisse. We also thank A. Shiryaev and C. Stricker for discussions on the topic. Part of the research was done while the first named author was visiting Tokyo University. Discussions with Prof. S. Kusuoka and Mrs. I. Nagayama greatly contributed to the paper.

Typeset by $\mathcal{AMS}$-TEX

1. Introduction

Roughly speaking, a passport option allows the holder, against paying a premium at the beginning of the contract, to take, during a predetermined time interval, positions in an underlying asset. These positions can be long or short, but are bounded. If at the end of the contract (maturity), the holder made a benefit, she can keep it. If on the contrary the holder made a loss, than he does not have to pay for these losses. That means the holder can keep the benefits but is not liable for the losses. The passport option is in some sense a generalisation of the American option. In the latter the holder can exercise the option only once. In the case of the passport option, the holder can "exercise" the option many times. Passport options were introduced by Bankers Trust, see [HLP].

There are different problems related to the passport option. First of all, there is the pricing problem. Since the holder can change the position many times, the price is given through an optimisation problem. The hedging problems are, at least in the complete market case, easily solved by standard methods. We will mainly focus on the pricing problem.

As said, the pricing problem is the result of an optimisation problem. The first paper on the subject was by Hyer, Lipton-Lifschitz and Pugachevsky, [HLP]. They were using methods from control theory in its relation with PDE's. The paper by Andersen, Andreasen and Brotherton-Ratcliffe, [ABB] treats the case of geometric Brownian motion through the use of stochastic control theory. Both papers calculate the price of the passport option as a solution to a PDE. The closed form solution however was not interpreted as an integral over a known distribution. In a series of talks between January 98 and January 99 (in Zurich, Tokyo, New York, Hong Kong, Toronto) the first named author presented the price calculation as an easy consequence of Skorohod's lemma and the use of local time. The idea was independently developed by Vicky Henderson and Hobson, [HH], and was then used to treat the more general Markov case. Numerical work in the so-called non symmetric case was developed by Nagayama, [Na]. She also treated the symmetric case using stochastic calculus and made a careful analysis of the smoothness of the value function. The discretisation procedure used in [Na] is different from ours as will be pointed out in section 7. A more recent paper on passport options is Shreve and Večeř, [SV].

The present paper picks up earlier results, but goes further in two directions. One is the relation with general martingale theory, the other direction deals with geometric Brownian motion, but in the presence of interest rate. This case is handled through time transforms in order to bring it back to a hitting time problem for Bessel processes. The techniques are similar to the ones developed by Geman and Yor and used to price Asian options, [GY] and [Y].

We do not handle the non-symmetric case. The optimal solution as calculated by Nagayama in [Na], shows that the switching boundary of the optimal strategy is non-trivial and its interpretation as a known curve is still open.

The paper is divided into several sections. This section will introduce some (standard) notation. Section 2 gives a description of different contracts and states the pricing problem as an optimisation problem. Section 3 relates the finiteness of the price to a characterisation problem of $\mathcal{H}^1$ semi-martingales. In section 4, the pricing problem is reduced to the calculation of the expected value of the one sided maximal function. This section is based on the relation between local time and the maximum functions. The basic ingredient is Skorohod's lemma. Section 5 is quite technical and mainly shows that for continuous martingales there is equality between two norms. This **equality** is related to the Davis **inequality** for $\mathcal{H}^1$ martingales. Section 6 deals with the discrete time optimisation problem for the geometric Brownian motion. Although stated in elementary terms, the proof of the main result is quite technical. Section 7 deals with the continuous time optimisation for the geometric Brownian motion. Here we pay attention to the non-existence of an optimal strategy. This non-existence is related to the non-existence of a strong solution for Tanaka's equation and to the difference between the filtrations generated by a Brownian motion B and its absolute value $|B|$. We also quickly discuss the relation with other approaches. Section 8 then treats the generalisation when interest rate is present. Although basically the same as the easy case treated in section 7, the solution requires more advanced technology. The basic ingredient is the fact that the geometric Brownian motion is a time transform of a Bessel process. To keep the paper as self contained as possible, we give full proofs of the intermediate results on Bessel processes. The main ingredient is the characterisation of the distribution of the hitting time of a Bessel process with a square root boundary.

The notation we use is standard. The structure $\left(\Omega, \mathcal{F} = (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ denotes a filtered probability space. The final sigma-algebra, if ever needed, is $\mathcal{F}_\infty = \bigvee_{t\geq 0} \mathcal{F}_t$. The filtration is supposed to satisfy the usual conditions, i.e. it is right continuous and $\mathcal{F}_0$ contains all the null sets of $\mathcal{F}_\infty$. Processes are defined on the time set $[0, t_0]$. The horizon t_0 is finite although we will never explicitly use this. Sometimes we will need Brownian motion. In that case it is not necessarily assumed that the filtration is generated by this Brownian motion. If so, we will explicitly say this. If B is a Brownian motion with respect to $\mathcal{F}$, then the geometric Brownian motion is the process defined as $S_t = \exp(B_t - t/2)$, i.e. the stochastic exponential of B. In applications to mathematical finance, this exponential plays a fundamental role. In order to simplify the notation, we have assumed that the price volatility is equal to 1. This can be achieved easily through an elementary time-transform. We invite the reader to apply the necessary changes.

The basic reference for martingale problems, Brownian motion and Bessel processes is Revuz-Yor, [RY].

2. Different contracts

There are many versions of a passport option. We just give a few examples. Although they differ, in at first sight only minor details, their price calculation can be completely different. The explicit analytic solution can (for the moment) only be given in some very specific cases. The underlying asset has a price evolution denoted by the semi-martingale S^0. The bank account is supposed to pay an interest rate given by the process r. The underlying S^0 pays out dividends at the rate f. In many applications, such as options on indices or the insurance of positions taken in mutual funds, the process f is identically zero. Both r and f are supposed to be sufficiently regular in order for the integrals $\int_0^{t_0} r_t\, dt$ and $\int_0^{t_0} f_t\, dt$ to exist a.s.. These assumptions do not play a special role in this paper, so we do not comment on them, the reader may well assume that both processes r and f are continuous and nonnegative. More important for us is the assumption of no arbitrage on the process S^0. From the general theory it follows that in order to be economically feasible, the discounted process, defined as

$$S_t = \exp\left(-\int_0^t (r_u - f_u)\, du\right) S_t^0,$$

should possess a local-martingale measure. For simplicity we already suppose the original measure $\mathbb{P}$ to be such that the process S is a local-martingale. For precise conditions on the existence of local-martingale measures and in the most general context, sigma-martingale measures, we refer the reader to [DS94] and [DS98]. We are now in a position to give some examples of passport options. Before doing so, we need one more notation. For each $t \in [0, t_0]$, we suppose that there is a set $Q_t \subset \mathbb{R}$ that describes the positions an investor can take in the asset S^0 at time t. This set should change in a measurable way with respect to t, more precisely:

$$\mathcal{Q} = \{(t, \omega, x) \mid x \in Q(t, \omega)\} \in \mathcal{P} \otimes \mathcal{R},$$

where $\mathcal{P}$ is the class of predictable sets and $\mathcal{R}$ is the class of Borel subsets of $\mathbb{R}$. Of course an investor should be able to take other positions as well, but these are then not covered by the passport option. A typical example would be when for all ω and all t we have that $Q_t = [-1, 1]$, or a little bit more general when $Q_t = [a, b]$ where $a < b$. We now give some examples:

(1) The investor receives dividends from the asset and receives interests on the bank account. The reference portfolio X^0 and its discounted value X_t defined as $X_t = \exp\left(-\int_0^t r_u\, du\right) X_t^0$, are then described as

$$\begin{aligned} q_t &\in Q_t \\ dX_t^0 &= q_t\, dS_t^0 + q_t S_t^0 f_t\, dt + \left(X_t^0 - q_t S_t^0\right) r_t\, dt \\ dX_t &= q_t \exp\left(-\int_0^t f_u\, du\right) dS_t. \end{aligned}$$

(2) The reference portfolio is described as if the investor would not collect the interest rate, but would collect the dividends. Such a passport option is probably not traded. In this case we find

$$\begin{aligned} q_t &\in Q_t \\ dX_t^0 &= q_t\, dS_t^0 + q_t S_t^0 f_t\, dt \\ dX_t &= q_t \exp\left(-\int_0^t f_u\, du\right) dS_t - r_t \left(X_t - q_t S_t \exp\left(-\int_0^t f_u\, du\right)\right) dt. \end{aligned}$$

(3) Similar to the previous one but this time the investor is entitled to the interest rate but not to the dividends. Contrary to the previous example this passport option seems to be traded. The description is:

$$\begin{aligned} q_t &\in Q_t \\ dX_t^0 &= q_t\, dS_t^0 + \left(X_t^0 - q_t S_t^0\right) r_t\, dt \\ dX_t &= q_t \exp\left(-\int_0^t f_u\, du\right) dS_t - q_t S_t \exp\left(-\int_0^t f_u\, du\right) f_t\, dt \\ &= q_t\, d\left(\exp\left(-\int_0^t f_u\, du\right) S_t\right). \end{aligned}$$

(4) Similar example, but this time there are no dividends and no interest rate. Also this version seems to exist only in theory.

$$\begin{aligned} q_t &\in Q_t \\ dX_t^0 &= q_t\, dS_t^0 \\ dX_t &= q_t\, d\left(\exp\left(\int_0^t r_u\, du\right)\exp\left(-\int_0^t f_u\, du\right) S_t\right). \end{aligned}$$

(5) If we suppose that dividends are reinvested in the asset, then it makes sense to adapt the bounds on the position, to this situation. This reduces to a change of the set Q_t in the following way. If the investor reinvests the dividends in the asset, then at time t he needs $\exp\left(\int_0^t f_u\, du\right)$ copies in order to obtain the same position as compared to a situation where one copy is held and where the asset would not pay out any dividends, but would itself reinvest these in the "world economy". In such a case it makes sense to replace the condition $q_t \in Q_t$ by the condition $q_t \in \exp\left(\int_0^t f_u\, du\right) Q_t$. Compared to example 1 above, this gives, the easier to handle

$$\begin{aligned} q_t &\in Q_t \\ dX_t &= q_t\, dS_t. \end{aligned}$$

(6) The investor can only rebalance its portfolio a limited number of times.
(7) The investor can only rebalance its portfolio once a day/week/month.
(8) the interest rate can be different when the portfolio is negative or when it is positive.

Needless to say that all these restrictions give rise to different problems in the calculation of the option price. In order to avoid more problems, similar to the calculation of an American option in an incomplete market, we make the following, loosely stated, assumptions.

Assumptions. *The market is supposed to satisfy the following properties*

(1) *The local martingale S is continuous.*
(2) *There is only one local martingale measure, e.g. the market is complete. We assume, for notational ease, that this measure is the given measure $\mathbb{P}$.*

In order to prepare for more general applications we will give some theorems that are valid for not necessarily continuous martingales. The appropriate assumptions will then clearly be stated and the notation will be adapted.
The price of a passport option can now be defined mathematically as follows. If the reference portfolio $X^0(q)$ and the discounted value $X(q)$, with starting point x_0, are defined as being dependent on the strategy q, then we are interested in the quantity

$$\sup\left\{\mathbf{E}_{\mathbb{P}}\left[(x_0 + X_{t_0}(q))^+\right] \mid q_t \in Q_t\right\}.$$

As the reader can check, this situation covers the examples above. For instance the case 6 is given by:

$$\sup\left\{\mathbf{E}_{\mathbb{P}}\left[\left(x_0 + \int_0^{t_0} q_u\, dS_u\right)^+\right] \mid q_t \in Q_t\right\}.$$

A special case is then

$$\sup\left\{\mathbf{E}_{\mathbb{P}}\left[\left(x_0+\int_0^{t_0} q_u\, dS_u\right)^+\right] \mid |q_t|\le 1\right\}.$$

There is a close relationship with the theory of $\mathcal{H}^1$ martingales and as we will see, if S is supposed to be the stochastic exponential of Brownian motion, then the quantity above can be calculated easily.

3. The relation with $\mathcal{H}^1$ Semi-Martingales

As seen in the examples, in the traded cases the reference portfolio is of the form

$$dX_t = q_t\, d\left(M_t A_t\right),$$

where M was a (local) martingale and where A was a process of finite variation. The following theorem describes under which conditions we can expect the passport option to have a finite price.

Theorem 3.1. *Suppose that $Z = MA$ is a (not necessarily continuous) semi-martingale where M is a local martingale and where A is a predictable process of finite variation. If we define*

$$\|Z\|_Q = \sup\left\{\mathbf{E}_{\mathbb{P}}\left[(q\cdot Z)^+_{t_0}\right] \mid q \text{ predictable and } |q|\le 1\right\}$$

and recall the definition of the $\mathcal{H}^1$ norm

$$\|Z\|_{\mathcal{H}^1} = \mathbf{E}_{\mathbb{P}}\left[\left(\int_0^{t_0} A_u^2\, d[M,M]_u\right)^{1/2} + \int_0^{t_0} |M_{u-}|\, |dA_u|\right],$$

then

$$c\,\|Z\|_{\mathcal{H}^1} \le \|Z\|_Q \le C\,\|Z\|_{\mathcal{H}^1},$$

where c and C are two universal constants.
The condition $\|Z\|_Q<\infty$ is equivalent to the condition (Var means total variation):

$$\mathbf{E}_{\mathbb{P}}\left[\sqrt{\int_0^{t_0} A_u^2\, d[M,M]_u}\right] < \infty \text{ and } \mathbf{E}_{\mathbb{P}}\left[\int_0^{t_0} |M_{u-}|\,|dA_u|\right] < \infty$$

In case M is a nonnegative, uniformly integrable martingale the latter requirement can be rewritten as $\mathbf{E}\left[M_{t_0} Var(A)_{t_0}\right] < \infty$.

Proof. Using Itô-calculus, the multiplicative Doob-Meyer decomposition can be transformed into an additive Doob-Meyer decomposition:

$$d(MA)_t = A_t\, dM_t + M_{t-}\, dA_t.$$

The theorem can now be proved along the same lines as the development of the $\mathcal{H}^p$ theory for semi-martingales, see [DM] and [Pr]. We prefer to include a proof

since the translation is not always that easy. The existence of the two universal constants will follow from the rest of the proof.
Since M is a local martingale and A is predictable, we may localise by stopping times T, such that M^T is an $\mathcal{H}^1$ martingale and A^T is a process of bounded variation. We then find that for each predictable process q, such that $|q| \leq 1$:

$$\begin{aligned}\mathbf{E}\left[\int_0^T q_u M_{u-}\, dA_u\right] &= \mathbf{E}\left[\int_0^T q_u\, dZ_u\right] \\ &\leq K = \sup\left\{\mathbf{E}_{\mathbb{P}}\left[(q\cdot Z)_{t_0}^+\right] \mid q \text{ predictable and } |q|\leq 1\right\} < \infty.\end{aligned}$$

This also means that for all q, predictable and bounded by 1 we have:

$$\mathbf{E}\left[M_T \int_0^T q_u\, dA_u\right] = \mathbf{E}\left[\int_0^T M_{u-} q_u\, dA_u\right] \leq K.$$

In particular we may take q so that we get $\int_0^T q_u M_{u-}\, dA_u = \int_0^T |M_{u-}|\, |dA_u|$. This yields

$$\mathbf{E}\left[\int_0^T |M_{u-}|\, |dA_u|\right] \leq K.$$

We then also find that for q predictable and bounded by 1:

$$\mathbf{E}\left[\left(\int_0^T q_u A_u\, dM_u\right)^+\right] \leq \mathbf{E}\left[(q\cdot Z)^+\right] + \mathbf{E}\left[\int_0^T |M_{u-}|\, |dA_u|\right] \leq 2K.$$

Of course this yields that

$$\mathbf{E}\left[\left|\int_0^T q_u A_u\, dM_u\right|\right] \leq 4K.$$

This implies that the martingale $A \cdot M$ is in $\mathcal{H}^1$, proving the first item of the theorem. A simple passage to the limit allows us to get rid of the localisation. The last statement follows easily since, by the predictability of A and hence of $Var(A)$:

$$\mathbf{E}\left[\int_0^{t_0} M_{u-}|dA_u|\right] = \mathbf{E}\left[M_{t_0}\int_0^{t_0} |dA_u|\right] = \mathbf{E}\left[M_{t_0} Var(A)_{t_0}\right].$$

□

Remark. The previous theorem shows that in order for the passport option to be meaningful, we have to require that the discounted price process S is in $\mathcal{H}^1$.

4. An Application of Skorohod's lemma.

In this section we will use the following lemma, due to Skorohod, see [RY], Chap VI.

Lemma 4.1. *If $s\colon \mathbb{R}_+ \to \mathbb{R}$ is a continuous function such that $s(0) \geq 0$, then we can write $z = s + l$ where*

1. *The function z is nonnegative and continuous.*
2. *The continuous function l is increasing, $l(0) = 0$ and the measure dl is supported by the (closed) set $\{z = 0\}$.*

Furthermore this decomposition is unique and the function l is given by

$$l(t) = \sup_{u \leq t} \left(-\min\left(0, s(u)\right)\right) = \sup_{u \leq t} s(u)^-.$$

There are many applications of this lemma to problems related to local time. The reader can check [RY] for details. In the following theorem as well as in the rest of the paper we will frequently use the notation Z^* for the *one sided* maximal function of a stochastic process Z, i.e.:

$$Z^*_t = \sup_{0 \leq u \leq t} Z_u.$$

Theorem 4.2. *If S is a continuous semi-martingale, if the process X satisfies $X_0 \geq 0$ and*

$$dX_t = -\operatorname{sign}(X_t) dS_t,$$

then

$$|X_t| = \sup_{u \leq t} \left(S_u - S_0 - X_0\right)^+ + X_0 + S_0 - S_t.$$

For $X_0 = 0$ this simplifies into

$$|X_t| = \sup_{u \leq t} S_u - S_t = S^*_t - S_0.$$

If moreover the process S is a uniformly integrable martingale and $X_0 = 0$, then

$$\mathbf{E}\left[|X_t|\right] = \mathbf{E}\left[S^*_t\right] - S_0.$$

Proof. The proof follows the proof of Lévy's theorem, see [RY] p. 230. Tanaka's formula gives us that

$$d|X|_t = -dS_t + dL_t(X),$$

where L is the local time of X at 0. It follows that

$$|X|_t = X_0 - (S_t - S_0) + L_t(X).$$

Because $X_0 \geq 0$, the process $S - S_0 + X_0$ starts at a nonnegative value. Furthermore the process L is increasing, satisfies $L_0 = 0$ and the measure dL is supported by the set $\{X = 0\}$. Skorohod's lemma tells us that

$$L_t = \sup_{u \leq t} \left(X_0 - S_u + S_0\right)^- = \sup_{u \leq t} \left(S_u - S_0 - X_0\right)^+.$$

The last statement of the theorem is obvious since $\mathbf{E}\left[S_t - S_0\right] = 0$. $\square$

5. An equality for the maximum of continuous martingales.

The aim of this section is to prove the following theorem:

Theorem 5.1. *If M is a continuous $\mathcal{H}^1$-martingale on $[0,t_0]$, if $M_0=0$, then*

$$\sup\left\{\mathbf{E}\left[(q\cdot M)^*_{t_0}\right] \mid q \text{ predictable and } |q|\le 1\right\} = \sup\left\{\mathbf{E}\left[|(q\cdot M)_{t_0}|\right] \mid q \text{ predictable and } |q|\le 1\right\}.$$

If the local martingale M is not in $\mathcal{H}^1$, then the equality remains valid in the sense that both quantities are $+\infty$.

Proof. We first deal with the easy case, i.e. where M is a local martingale that is not in $\mathcal{H}^1$. Clearly we have that

$$\sup_{u\le t_0}|M_u| \le M^*_{t_0} + (-M)^*_{t_0}.$$

By the Davis' inequality, see [RY] or [Pr], the two norms

$$\mathbf{E}\left[\sup_{u\le t_0}|M_t|\right] \text{ and } \sup\left\{\mathbf{E}\left[|(q\cdot M)_{t_0}|\right] \mid q \text{ predictable and } |q|\le 1\right\},$$

are equivalent. Hence we find that

$$\sup\left\{\mathbf{E}\left[|(q\cdot M)_{t_0}|\right] \mid q \text{ predictable and } |q|\le 1\right\}$$

and

$$\sup\left\{\mathbf{E}\left[(q\cdot M)^*_{t_0}\right] \mid q \text{ predictable and } |q|\le 1\right\}$$

are at the same time finite or infinite. The amazing thing is that, in case both quantities are finite, they are equal. This is more precise than what the Davis' inequality shows.

The usual convexity arguments allow us to restrict the analysis to predictable processes q such that $|q|=1$. This is done as follows. The unit ball of the Banach space $L^\infty(\Omega\times[0,t_0],\mathcal{P},d\mathbb{P}\otimes d\langle M,M\rangle)$, seen as the dual of the Banach space $L^1(\Omega\times[0,t_0],\mathcal{P},d\mathbb{P}\otimes d\langle M,M\rangle)$, is the set of all predictable processes, bounded by 1. The extreme points are the processes q such that $|q|=1$. By weak* compactness and the Krein-Milman theorem, the unit ball is also the weak*-closed convex hull of its extreme points, and hence the convex hull of the extreme points is dense in the unit ball for the convergence in measure. The latter follows from the fact that on the unit ball, the topology of convergence in measure is precisely the Mackey topology of the dual pair (L^1,L^∞). Hence the closed convex hull for the weak* topology coincides with the closed convex hull for the convergence in measure. The rest now follows from the dominated convergence theorem for stochastic integrals. We get that

$$\sup\left\{\mathbf{E}\left[(q\cdot M)^*_{t_0}\right] \mid q \text{ predictable and } |q|\le 1\right\} = \sup\left\{\mathbf{E}\left[(q\cdot M)^*_{t_0}\right] \mid q \text{ predictable and } |q|= 1\right\}$$

as well as

$$\sup\left\{\mathbf{E}\left[|(q\cdot M)_{t_0}|\right] \mid q \text{ predictable and } |q|\le 1\right\} = \sup\left\{\mathbf{E}\left[|(q\cdot M)_{t_0}|\right] \mid q \text{ predictable and } |q|= 1\right\}.$$

One inequality is almost trivial and follows from Tanaka's formula and Skorohod's lemma. Indeed for q predictable and of modulus 1 we obtain that

$$d|q \cdot M|_t = \operatorname{sign}((q \cdot M)_t)\, q_t\, dM_t + dL_t,$$

where L is the local time of $q \cdot M$ at 0. Skorohod's lemma tells us that

$$L_t = \sup_{u \le t}(h \cdot M)_u \text{ where } h_u = -\operatorname{sign}((q \cdot M)_u)\, q_u.$$

It follows that

$$\mathbf{E}\left[|(q \cdot M)_{t_0}|\right] = \mathbf{E}\left[L_{t_0}\right] \le \sup\left\{\mathbf{E}\left[(p \cdot M)^*_{t_0}\right] \mid p \text{ predictable and } |p| \le 1\right\}.$$

The other inequality is less trivial. For given h, predictable and $|h| = 1$, we put $N = h \cdot M$. One way to prove the remaining inequality could consist in finding a solution to the equation

$$dX_u = -\operatorname{sign}(X_u)\, dN_u.$$

In the case where N (or M) is a Brownian motion, such equations have in general only weak solutions. In the case of general continuous martingales, the concept of weak solution is not easily understood. Our proof uses discrete time approximations. It has an interest in itself. In accordance with stochastic practice, we put $\operatorname{sign}(0) = -1$.

Lemma 5.2. *Let N be a continuous $\mathcal{H}^1$-martingale, defined on the time interval $[0, t_0]$ and starting at 0. For each n let there be given a finite sequence of stopping times,*

$$0 = \tau^n_0 \le \tau^n_1 \le \ldots \le \tau^n_{K_n} = t_0.$$

Let the martingale X^n be defined as the solution of

$$X^n_0 = 0 \text{ and } dX^n_t = -\operatorname{sign}(X_{\tau^n_k})\, dN_t \quad \text{for } \tau^n_k < t \le \tau^n_{k+1}.$$

If $\max_{1 \le k \le K_n}\left(\tau^n_{k+1} - \tau^n_k\right)$ tends to zero in probability, then $\mathbf{E}[|X^n_{t_0}|]$ tends to the quantity $\mathbf{E}[\sup_{t \le t_0} N_t]$. More precisely the predictable process α defined on $]\!]\tau^n_k, \tau^n_{k+1}]\!]$ as $\alpha_t = \operatorname{sign}\left(X_{\tau^n_k}\right)\operatorname{sign}\left(X_t\right)$ tends to 1 on $[0, t_0] \times \Omega$.

Proof of lemma 5.2. We first show how the statement on the sequence α^n leads to the other result. We introduce the σ-finite measure μ on the $\mathcal{R} \otimes \mathcal{F}_{t_0}$ measurable sets P of $[0, t_0] \times \Omega$ as follows:

$$\mu(P) = \mathbf{E}_{\mathbb{P}}\left[\int_{[0,t_0]} \mathbf{1}_P\, d\langle N, N\rangle\right].$$

What we claim is that on sets of finite μ-measure, the sequence of predictable processes α^n tends to 1 in μ-measure , i.e. for each predictable set P such that $\mu(P) < \infty$ we have that $\mu\left(\{\alpha^n \ne 1\} \cap P\right) \to 0$. Now the Itô-Tanaka formula gives that

$$d|X^n_t| = -\alpha^n_t\, dN_t + dL^n_t,$$

where L^n is the local time at zero of the process X^n. Skorohod's lemma implies that $L^n_t = \sup_{0\le s\le t}(\alpha^n \cdot N)_s$. But the convergence of α^n implies that

$$\int_{[0,t_0]} (\alpha^n_t - 1)^2 \, d\langle N, N\rangle \to 0,$$

in probability $\mathbb{P}$. Since N is an $\mathcal{H}^1$ martingale we get that $\alpha^n \cdot N$ tends to N in $\mathcal{H}^1$ and hence we get that, in L^1:

$$\sup_{t\le t_0}(\alpha^n \cdot N)_t \to \sup_{t\le t_0} N_t.$$

But then we also have that

$$\mathbf{E}\left[\sup_{t\le t_0} N_t\right] = \lim_n \mathbf{E}\left[\sup_{t\le t_0}(\alpha^n \cdot N)_t\right] = \lim_n \mathbf{E}[|X^n_{t_0}|].$$

So we only have to prove the statement about the sequence α^n. We will do this through a time-transform of the martingale N into a Brownian motion. To have the transform well defined, we continue the martingale N, beyond t_0 with an independent Brownian motion. This is standard as can be seen from [RY] p. 174. In order to do this we first time-transform, in case $t_0 = \infty$, the interval $[0, t_0]$ into $[0, 1]$. The filtration is extended in the obvious way, see RY, p 174. The extension will still be denoted by N. The DDS time changes C_u are now defined as

$$C_u = \inf\{t \mid \langle N, N\rangle_t \ge u\}.$$

Because we reduced the problem to $t_0 < \infty$ and continued N with an independent Brownian motion, these stopping times are finite almost surely. The process $\beta_u = N_{C_u}$ defines a Brownian motion with respect to the filtration $(\mathcal{F}_{C_u})_{u\ge 0}$. In particular for $t \ge u$ we have that the process $(\beta_t - \beta_u)_{t\ge u}$ is independent of $\mathcal{F}_{C_u}$. We also extend, in the same way the measure μ to the sigma-algebra $\mathcal{R} \otimes \mathcal{F}_\infty$.
We next fix $\delta > 0$ as well as $A > 0$ and we will show that there is an absolute constant c such that for all n big enough we have

$$\mu\left(\{\alpha^n \ne 1\} \cap [\![0, C_A]\!]\right) \le \left(A + c\sqrt{A}\right)\sqrt{\delta}.$$

This will then end the proof of the lemma.
We observe that by continuity of N we have that

$$\mathbb{P}\left[\max_k \left(\langle N, N\rangle_{\tau^n_{k+1}} - \langle N, N\rangle_{\tau^n_k}\right) > \delta\right] \to 0.$$

So for n big enough this quantity will be smaller than δ. This is the only bound on n we need. So from now on we assume that n is big enough and fixed. For notational ease we also define new families of stopping times. First we extend the family of stopping times $\left(\tau^n_j\right)_{1\le j\le K_n}$ by defining τ^n_j for $j > K_n$ as follows. The stopping time $\tau^n_{K_n+1}$ is the first time t, after t_0, for which $\langle N, N\rangle_t - \langle N, N\rangle_{t_0} \ge \delta$. Recursively we define τ^n_{j+1}, $j \ge K_n + 1$ as the first time $t \ge \tau^n_j$ we have $\langle N, N\rangle_t - \langle N, N\rangle_{\tau^n_j} \ge \delta$. To keep notation consistent and for notational ease we also extend the processes

X^n beyond the interval $[0, t_0]$. This is done using the same differential equation. A very important fact is that the DDS-time change of the processes X^n and N are the same, namely the family $(C_u)_{0\leq u}$. In particular the variables $X^n_{C_u}$ are gaussian with variance u.
We now define the sequence of stopping times V_k (they depend also on n, but we drop the index for notational ease) as follows.

$$V_0 = 0$$
$$V_1 = \tau^n_j,$$

where j is the first index for which $\tau_j \geq C_\delta$. Having defined V_k we define the next stopping time V_{k+1} as:

$$V_{k+1} = \tau^n_j,$$

where j is the first index such that $\tau^n_j \geq V_k$ and such that $\tau_j \geq C_{(k+1)\delta}$.
We observe that on the set $\left\{\max_{1\leq j\leq K_n}\left(\tau^n_j - \tau^n_{j-1}\right) \leq \delta\right\} = \left\{\max_{1\leq j}\left(\tau^n_j - \tau^n_{j-1}\right) \leq \delta\right\}$ we have that for all k: $C_{k\delta} \leq V_k \leq C_{(k+1)\delta}$ and hence $\max_k\left(V_{k+1} - V_k\right) \leq 2\delta$. Furthermore for all $\omega \in \Omega$ the number of elements in the set $\{k \mid V_k(\omega) \leq C_A(\omega)\}$ is bounded by $(A/\delta + 1)$.
The next step is to find sets that are bigger than $\{\alpha^n \neq 1\} \cap [\![0, C_A]\!]$ and for which the μ-measure can be calculated easily. This is done as follows:

$$\{\alpha^n \neq 1\} \cap [\![0, C_A]\!] = \{(t,\omega) \mid 0 \leq t \leq C_A; \alpha^n_t(\omega) \neq 1\}$$

$$\subset \bigcup_{j\geq 0}\left\{(t,\omega) \mid \begin{array}{l} 0 \leq t \leq C_A(\omega); \tau^n_j(\omega) \leq t \leq \tau^n_{j+1}(\omega); \\ \mathrm{sign}(X^n_t(\omega)) \neq \mathrm{sign}(X^n_{\tau^n_j}(\omega) \end{array}\right\}$$

$$\subset \bigcup_{j\geq 0}\left\{(t,\omega) \mid \begin{array}{l} 0 \leq t \leq C_A(\omega); \tau^n_j(\omega) \leq t \leq \tau^n_{j+1}(\omega); \\ \mathrm{sign}(X^n_s(\omega)) \text{ not constant on } [\tau^n_j(\omega), \tau^n_{j+1}(\omega)] \end{array}\right\}$$

$$\subset \bigcup_{k\geq 0}\left\{(t,\omega) \mid \begin{array}{l} 0 \leq t \leq C_A(\omega); V_k(\omega) \leq t \leq V_{k+1}(\omega); \\ \mathrm{sign}(X^n_s(\omega)) \text{ not constant on } [V_k(\omega), V_{k+1}(\omega)] \end{array}\right\}$$

$$\subset \left\{(t,\omega) \mid \max_{0\leq j\leq K_n-1}(\tau^n_{j+1} - \tau^n_j)(\omega) > \delta; 0 \leq t \leq C_A(\omega)\right\}$$
$$\bigcup \bigcup_{0\leq k\leq \frac{A}{\delta}}\left\{(t,\omega) \mid \begin{array}{l} C_{k\delta}(\omega) \leq t \leq C_{(k+2)\delta}(\omega); \\ \mathrm{sign}(X^n_s(\omega)) \text{ not constant on } [C_{k\delta}(\omega), C_{(k+1)\delta}(\omega)] \end{array}\right\}.$$

The μ measure of the first set is bounded as follows:

$$\mu\left(\left\{(t,\omega) \mid \max_{0\leq j\leq K_n-1}(\tau^n_{j+1} - \tau^n_j)(\omega) > \delta; 0 \leq t \leq C_A(\omega)\right\}\right) \leq A\delta.$$

For each $k = 0, \dots, \frac{A}{\delta}$ we have

$$\mu\left(\left\{(t,\omega) \mid \begin{array}{l} C_{k\delta}(\omega) \le t \le C_{(k+2)\delta}(\omega); \\ \operatorname{sign}(X^n_s(\omega)) \text{ not constant on } [C_{k\delta}(\omega), C_{(k+1)\delta}(\omega)] \end{array}\right\}\right)$$

$$\le \mu\left(\left\{(t,\omega) \mid \begin{array}{c} C_{k\delta}(\omega) \le t \le C_{(k+2)\delta}(\omega); \\ \sup_{0\le s\le 2\delta} \left|X^n_{s+C_{k\delta}} - X^n_{C_{k\delta}}\right| > \left|X^n_{C_{k\delta}}\right| \end{array}\right\}\right)$$

$$\le 2\delta \mathbb{P}\left[\sup_{0\le s\le 2\delta} \left|X^n_{s+C_{k\delta}} - X^n_{C_{k\delta}}\right| > \left|X^n_{C_{k\delta}}\right|\right].$$

The latter probability can be calculated using the independence properties of the Brownian motion together with an estimate on the probability that a Brownian motion crosses a level within a time interval of length 2δ, see [RY] page 70, Exercise 3.14 as well as p.320, exercise 1.21. We get that (c is a constant that can change from one line to another)

$$\begin{aligned}
&\mathbb{P}\left[\sup_{0\le s\le 2\delta} \left|X^n_{s+C_{k\delta}} - X^n_{C_{k\delta}}\right| > \left|X^n_{C_{k\delta}}\right|\right] \\
&\le c\,\mathbf{E}\left[\exp\left(-\frac{\left(X^n_{C_{k\delta}}\right)^2}{4\delta}\right)\right] \\
&\le c\int_{-\infty}^{+\infty} e^{-\frac{y^2}{4\delta}} \frac{1}{\sqrt{2\pi k\delta}} e^{-\frac{y^2}{2k\delta}}\, dy \\
&\le c\int_{-\infty}^{+\infty} e^{-\frac{u^2}{2} - \frac{u^2 k\delta}{4}}\, du \\
&\le \frac{c}{\sqrt{k+1}}.
\end{aligned}$$

Putting together all the estimates gives us that

$$\begin{aligned}
&\mu\left(\{\alpha^n \ne 1\} \cap [\![0, C_A]\!]\right) \\
&\le A\delta + c\delta \sum_{0\le k\le \frac{A}{\delta}} \frac{1}{\sqrt{k+1}} \\
&\le A\delta + c\delta\sqrt{\frac{A}{\delta}} \\
&\le (A + c\sqrt{A})\sqrt{\delta}.
\end{aligned}$$

This completes the proof of the lemma 5.2. $\square$

We now continue the proof of theorem 5.1. For given $|q| = 1$ predictable, we define $N = q\cdot M$ and apply the lemma 5.2 to the grid defined by $\left(\frac{jt_0}{2^n}\right)_{0\le j\le 2^n}$. With the notation of the lemma we get that

$$\mathbf{E}\left[\sup_{0\le t\le t_0} (q\cdot M)_t\right] = \lim \mathbf{E}\left[\left|X^n_{t_0}\right|\right] \le \sup\left\{\mathbf{E}\left[\left|(h\cdot M)_{t_0}\right|\right] \mid q \text{ predictable and } |h| \le 1\right\}.$$

This ends the proof of the theorem. $\square$

6. The Discrete Time Optimal Solution

In this section we suppose that the process $S_t = \exp\left(W_t - \frac{1}{2}t\right)$ is a geometric Brownian motion defined on some filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\leq t_0}, \mathbb{P}\right)$. We do not suppose that the filtration $\mathcal{F}$ is generated by the Brownian motion W. Suppose that a trader can only trade at given fixed dates $0 = t_0 \leq t_1 \leq t_2 \ldots \leq t_N = t_0$. He/she then wants to select a predictable strategy q, $|q| \leq 1$, that is constant on the intervals $]\!]t_k, t_{k+1}]\!]$ and that maximises, for given $x \in \mathbb{R}$, the quantity

$$\mathbf{E}_{\mathbb{P}}\left[\left|x + \sum_{0\leq k\leq N-1} q_{t_k}\left(S_{t_{k+1}} - S_{t_k}\right)\right|\right].$$

Theorem 6.1. *If the process X and the process q are defined as*

$$X_0 = x \in \mathbb{R} \quad \textit{and } q_0 = -1 \textit{ if } x = 0$$
$$X_t = X_{t_k} + q_t\left(S_t - S_{t_k}\right); q_t = -\operatorname{sign}(X_{t_k}) \quad \textit{for } t_k < t \leq t_{k+1},$$

Then q is the optimal strategy. This means that for every sequence of functions $(f_k)_{0\leq k\leq N-1}$, such that $|f_k| \leq 1$ and f_k being $\mathcal{F}_{t_k}$ measurable, we have

$$\mathbf{E}_{\mathbb{P}}\left[\left|x + \sum_{0\leq k\leq N-1} f_k\left(S_{t_{k+1}} - S_{t_k}\right)\right|\right] \leq \mathbf{E}_{\mathbb{P}}\left[\left|x + \sum_{0\leq k\leq N-1} q_{t_k}\left(S_{t_{k+1}} - S_{t_k}\right)\right|\right].$$

The proof of this statement is not obvious and will be divided into several lemmata. The idea is to use dynamic programming. That means we first try to calculate the optimal solution when $N = 1$ as well as the corresponding value function and then we proceed by backward recursion. The one time step case is easy and is solved in the following lemma.

Lemma 6.2. *For each $t \in \mathbb{R}_+$, $x \geq 0$ and $s \geq 0$ we have that*

$$\mathbf{E}\left[\left|x + s\left(S_t - 1\right)\right|\right] \leq \mathbf{E}\left[\left|x - s\left(S_t - 1\right)\right|\right]$$

Proof of lemma 6.2. We distinguish two cases.
case 1: $x \geq s$. This is the easy case since

$$\left|x + s\left(S_t - 1\right)\right| = x + s\left(S_t - 1\right),$$

and therefore $x = \mathbf{E}\left[\left|x + s\left(S_t - 1\right)\right|\right]$. However the random variable $x - s\left(S_t - 1\right)$ has mean x and may assume negative values. Therefore we have that

$$x < \mathbf{E}\left[\left|x - s\left(S_t - 1\right)\right|\right].$$

case 2: $0 \leq x < s$. This requires a better reasoning. Let us define the measure $\mathbb{Q}$

as $d\mathbb{Q} = S_t\, d\mathbb{P}$. We then have that

$$\begin{aligned}
&\int |x + s\,(S_t - 1)|\; d\mathbb{P} \\
&= \int \left|xS_t^{-1} + s - sS_t^{-1}\right| S_t\, d\mathbb{P} \\
&= \int \left|x - (s - x)\left(S_t^{-1} - 1\right)\right|\; d\mathbb{Q} \\
&\qquad \text{and since } S_t^{-1} \text{ under } \mathbb{Q} \text{ has the same law as } S_t \text{ under } \mathbb{P} \\
&= \int |x - (s - x)\,(S_t - 1)|\; d\mathbb{P} \\
&\leq \left(1 - \frac{x}{s}\right) \int |x - s\,(S_t - 1)|\; d\mathbb{P} + \frac{x}{s} \int x\, d\mathbb{P} \\
&\leq \left(1 - \frac{x}{s}\right) \int |x - s\,(S_t - 1)|\; d\mathbb{P} + \frac{x}{s} \int |x - s\,(S_t - 1)|\; d\mathbb{P} \\
&\leq \int |x - s\,(S_t - 1)|\; d\mathbb{P}.
\end{aligned}$$

□

Remark. Of course this lemma could have been proved by direct calculation, using the density of the lognormal distribution. We preferred to give a more structural proof.

We now recursively define functions $\Psi_k\colon \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R}_+$. These functions (so called value functions) are defined as

$$\Psi_0(x, s) = |x|$$

$$\Psi_{k+1}(x, s) = \max \begin{cases} \mathbf{E}\left[\Psi_k\left(x + s\left(\dfrac{S_{t_{N-k}}}{S_{t_{N-k-1}}} - 1\right), s\dfrac{S_{t_{N-k}}}{S_{t_{N-k-1}}}\right)\right] \\ \mathbf{E}\left[\Psi_k\left(x - s\left(\dfrac{S_{t_{N-k}}}{S_{t_{N-k-1}}} - 1\right), s\dfrac{S_{t_{N-k}}}{S_{t_{N-k-1}}}\right)\right] \end{cases}$$

In order to show that in the maximum above, the greater value is attained for $-\operatorname{sign}(x)$, we have to look for properties of Ψ_k. The relevant properties are listed in the following lemma.

Lemma 6.3. *The functions Ψ_k defined above satisfy the following properties*

(1) $\Psi_k(-x, s) = \Psi_k(x, s)$
(2) $\Psi_k(\lambda x, \lambda s) = \lambda \Psi_k(x, s)$ *for all* $\lambda > 0$
(3) *they are convex on* $\mathbb{R} \times \mathbb{R}_+$
(4) $\Psi_k(x, s) \geq |x|$
(5) $\lim_{x\to\infty} \frac{\Psi_k(x,s)}{|x|} = 1$
(6) $\Psi_k(x, s) \leq \Psi_k(0, s) + |x|$
(7) *If* $\Psi_k'(x, s)$ *denotes the left (or right) derivative of* Ψ *with respect to the first variable, then* $\lim_{x\to-\infty} \Psi_k'(x, s) = -1$ *and* $\lim_{x\to+\infty} \Psi_k'(x, s) = +1$
(8) $\lim_{s\to 0} \Psi_k(x, s) = |x|$ *uniformly on* $\mathbb{R}$.

Proof of lemma 6.3. The proof is done by induction on k. The reader can check that properties 1, 2 follow from the definition and by induction the convexity in property

3 can be verified by direct inspection. We now prove property 4 by induction. Let us fix a time $t \geq 0$ and have a look at the functions

$$\Psi_k^+(x,s) = \mathbf{E}\left[\Psi_{k-1}\left(x + s(S_t - 1), sS_t\right)\right]$$
$$\Psi_k^-(x,s) = \mathbf{E}\left[\Psi_{k-1}\left(x - s(S_t - 1), sS_t\right)\right].$$

An application of Jensen's inequality immediately yields that $\Psi_k^{\pm}(x,s) \geq \Psi_{k-1}(x,s) \geq |x|$, where the last step is the induction hypothesis. If we apply the above reasoning for $t = t_{N-k+1} - t_{N-k}$ and observe, as will be done several times below, that S_t has the same law as $\frac{S_{t_{N-k+1}}}{S_{t_{N-k}}}$, we get that $\Psi_k(x,s) = \max\left(\Psi_k^+(x,s), \Psi_k^-(x,s)\right) \geq |x|$. We now prove the remaining properties. Remark that they are obvious for $k = 0$. So we concentrate on the induction step. Let us start with property 6.

$$\begin{aligned}\Psi_k^+(x,s) &= \mathbf{E}\left[\Psi_{k-1}\left(x + s(S_t - 1), sS_t\right)\right] \\ &\leq \mathbf{E}\left[\Psi_{k-1}\left(0, sS_t\right) + |x| + s\left|S_t - 1\right|\right] \\ &\leq s\mathbf{E}\left[S_t\Psi_{k-1}(0,1) + |S_t - 1|\right] + |x|\end{aligned}$$

So it follows that there is a constant a such that $\Psi_k^+(x,s) \leq as + |x|$. This inequality, by the way, also implies (by induction) that $\Psi_k < +\infty$. But now convexity implies that for each n we have $\Psi_k^+(x,s) \leq \frac{n-1}{n}\Psi_k^+(0,s) + \frac{1}{n}\Psi(nx,s) \leq \frac{n-1}{n}\Psi_k^+(0,s) + \frac{1}{n}\left(as + n|x|\right)$. If we let n tend to infinity this gives $\Psi_k^+(x,s) \leq \Psi_k^+(0,s) + |x|$. The same reasoning applies to Ψ_k^- and hence the inequality remains true for Ψ_k. Property readily follows from properties 4 and 6. The last property 8 will follow from 6 and the property $\Psi_k(0,s) \to 0$ as $s \to 0$. The latter can be checked easily:

$$\Psi_k(0,s) = s\Psi_k(0,1) \to 0 \text{ since } \Psi_k(0,1) < +\infty.$$

□

Lemma 6.4. *If a function $\psi: \mathbb{R} \to \mathbb{R}_+$ satisfies the properties*

(1) *ψ is convex*
(2) $\psi(x) = \psi(-x)$
(3) *$\psi(x) \geq |x|$ and $\lim_{x\to+\infty} \frac{\psi(x)}{x} = 1$,*

then there exists a real number α as well as a probability measure λ on $\mathbb{R}_+$ such that

$$\psi(x) = \alpha + \int_{\mathbb{R}_+} \max(|x|, a)\, \lambda(da).$$

Proof of lemma 6.4. The proof of this is a slight adaptation of the representation theorem for convex functions, see e.g. [RY], appendix. For completeness we give a sketch. The convexity relation $\psi(y) - \psi(x) \geq \psi'_+(x)(y - x)$ yields the following. We fix x, divide by y and let y tend to $-\infty$. This gives $-1 \leq \psi'_+(x)$ for all x. Now we divide by x and let x tend to $-\infty$. This gives that $\lim \psi'_+(x) \leq -1$. I.e. $\lim_{x\to-\infty} \psi'_+(x) = -1$. The function $\psi(x) + x$ is decreasing for x tending to $-\infty$ and is bounded below by 0 (because of the hypothesis). Let the limit be a. All this allows us to write $\psi(x) = a - x + g(x)$, where g is convex, nonnegative, tends to 0 for $x \to -\infty$. Furthermore the measures ψ'' and g'' coincide. Write now

$g(x) = \int_{(-\infty,x]} g'_+(u)\,du$, substitute $g'_+(u) = \int_{(-\infty,u]} g''(dy)$, apply Fubini's theorem to get

$$\psi(x) = a - x + \int_{(-\infty,x]} \psi''(dy)(x-y) = \int_{\mathbb{R}} \psi''(dy)(x-y)^+.$$

An analysis of the behaviour of ψ near $+\infty$ yields a similar result and this gives the existence of a number b such that

$$\psi(x) = b + x + \int_{[x,+\infty)} \psi''(dy)(y-x) = \int_{\mathbb{R}} \psi''(dy)(y-x)^+.$$

Adding the two expressions yields the existence of a number c such that

$$\psi(x) = c + \frac{1}{2}\int_{\mathbb{R}} |x-y|\psi''(dy).$$

Until now we did not use the symmetry of the function ψ and the analysis could have been given for convex functions, having linear behaviour at $\pm\infty$. The symmetry of ψ now gives that

$$\psi(x) = c + \int_{\mathbb{R}} \frac{1}{2}\psi''(dy)\frac{|x-y|+|x+y|}{2}.$$

Furthermore $\frac{|x-y|+|x+y|}{2} = \max(|x|,|y|)$ and by symmetry

$$\psi(x) = c + \frac{1}{2}\int_{\mathbb{R}} \psi''(dy)\max(|x|,|y|) = c + \int_{\mathbb{R}_+\setminus\{0\}} \psi''(dy)\max(|x|,y) + \frac{1}{2}\psi''(\{0\})|x|.$$

Since $\psi''(\mathbb{R}) = \psi'(+\infty) - \psi'(-\infty) = 2$, we rewrite this as

$$\psi(x) = c + \int_{\mathbb{R}_+} \lambda(dy)\max(|x|,y),$$

where λ is a probability measure on $\mathbb{R}_+$. $\square$

Before we can solve the main technical difficulty in the optimisation problem, we need one more lemma, which is an easy application of elementary analysis. We do not aim for the most general form of this lemma.

Lemma 6.5. *Let $f\colon \mathbb{R}_+ \to \mathbb{R}$ be a continuous function that is everywhere differentiable on the open interval $]0,\infty[$. Suppose that*

1. $\lim_{x\to\infty} f(x) = 0$
2. *for z big enough we have that $f(z) \geq 0$*
3. $f(0) \geq 0$
4. *there is at most one $y_0 \in]0,\infty[$ such that $f'(y_0) = 0$*

then $f(x) \geq 0$ for all $x \geq 0$.

Proof of lemma 6.5. We extend the function f to the closed interval $[0,\infty]$ by putting $f(\infty) = 0$. By compactness there is at least one point z_0 where f attains its minimum, of course $f(z_0) \leq 0$. If z_0 is either 0 or ∞ then clearly $f(z) \geq f(z_0) \geq 0$ for all $z \in \mathbb{R}$. So we may suppose that $z_0 \in]0,\infty[$. Since f is differentiable we get that $f'(z_0) = 0$. Let now z_1 be a point where f attains its maximum on $[z_0,\infty)$.

Since $f(z) \geq 0$ for z big enough, we may suppose that $z_1 \neq \infty$. If $z_1 = z_0$, then clearly $f(z_0) = f(z_1) \geq 0$, implying that f is equal to zero on $[z_0, \infty]$, a contradiction to item 4. If on the contrary $z_1 \neq z_0$ then by differentiability of f we have that z_1 is a second point where $f'(z_1) = 0$, again a contradiction. □

Remark. A careful inspection of the proof shows that under the same hypothesis as in lemma 6.5, the extra assumptions that $f(z) > 0$ for all z big enough, yields that $f(z) > 0$ for all $z > 0$.

Lemma 6.6. *If* $\psi: \mathbb{R} \to \mathbb{R}_+$ *satisfies*

(1) ψ *is convex,*
(2) $\psi(x) = \psi(-x)$,
(3) $\psi(x) \geq |x|$ *and* $\lim_{x\to+\infty} \frac{\psi(x)}{x} = 1$,

then for all $x \geq 0$, $t \geq 0$, $s > 0$ *we have that*

$$\mathbf{E}\left[\psi\left((s-x)S_t - s\right)\right] \leq \mathbf{E}\left[\psi\left((s+x)S_t - s\right)\right]$$

Proof of lemma 6.6. Since there is equality if $x = 0$ we may, from now on, suppose that $x > 0$. As in the proof of lemma 6.2 we distinguish two cases.
case 1: $x \geq s$. The inequality is an easy consequence of convexity. Indeed

$$\begin{aligned}
\mathbf{E}\left[\psi((s-x)S_t - s)\right] &= \mathbf{E}\left[\psi((x-s)S_t + s)\right] \\
&= \mathbf{E}\left[\psi(x + (x-s)(S_t - 1))\right] \\
&\leq \frac{x-s}{x+s}\mathbf{E}\left[\psi(x + (x+s)(S_t - 1))\right] + \left(1 - \frac{x-s}{x+s}\right)\psi(x) \\
&\leq \mathbf{E}\left[\psi(x + (x+s)(S_t - 1))\right] \quad \text{by Jensen's inequality} \\
&\leq \mathbf{E}\left[\psi((x+s)S_t - s)\right]
\end{aligned}$$

case 2: $0 < x < s$. This is the non-trivial case. Of course, by the representation lemma 6.4, we only have to treat the case $\psi(x) = \max(|x|, a)$. We define two functions on $\mathbb{R}_+$:

$$\begin{aligned}
\phi(+, a) &= \mathbf{E}\left[\max\left(a, |(s+x)S_t - s|\right)\right] \\
\phi(-, a) &= \mathbf{E}\left[\max\left(a, |(s-x)S_t - s|\right)\right].
\end{aligned}$$

Clearly these two functions are convex and Lipschitz. For $a \geq s$ we furthermore have that

$$0 \leq \max\left(a, |(s+x)S_t - s|\right) - \max\left(a, |(s-x)S_t - s|\right) \leq 2xS_t \mathbf{1}_{S_t \geq \frac{a+s}{s+x}}.$$

This implies that for $a \geq s$ we have that $\phi(+, a) - \phi(-, a) \geq 0$ and by Lebesgue's dominated convergence theorem we get $\lim_{a\to+\infty} \phi(+, a) - \phi(-, a) = 0$. A straightforward application of Lebesgue's theorem again, allows us to calculate the derivative of $\phi(+, a) - \phi(-, a) \geq 0$ by passing under the integral sign. This yields:

$$\begin{aligned}
\frac{d\phi(+, a)}{da} &= \mathbb{P}\left[|(s+x)S_t - s| \leq a\right] \\
\frac{d\phi(-, a)}{da} &= \mathbb{P}\left[|(s-x)S_t - s| \leq a\right]
\end{aligned}$$

For $a \geq s$ we know that

$$\{|(s+x)S_t - s| \leq a\} \subsetneq \{|(s-x)S_t - s| \leq a\},$$

and hence we get that $\frac{d(\phi(+,a)-\phi(-,a))}{da} < 0$. This also means that the difference $\phi(+,a) - \phi(-,a)$ decreases to zero for $a \geq s$. For $a \leq s$ we proceed as follows.

$$\begin{aligned}
\{|(s+x)S_t - s| \leq a\} &= \{-a \leq (s+x)S_t - s \leq a\} \\
&= \left\{\frac{s-a}{s+x} \leq S_t \leq \frac{s+a}{s+x}\right\} \\
&= \left\{\frac{1}{2}t + \ln\left(\frac{s-a}{s+x}\right) \leq W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s+x}\right)\right\} \\
\{|(s-x)S_t - s| \leq a\} &= \{-a \leq (s-x)S_t - s \leq a\} \\
&= \left\{\frac{s-a}{s-x} \leq S_t \leq \frac{s+a}{s-x}\right\} \\
&= \left\{\frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right) \leq W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right)\right\}.
\end{aligned}$$

These equalities imply that

$$\mathbb{P}\left[|(s+x)S_t - s| \leq a\right] = \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s+x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s+x}\right)\right] \text{ and}$$

$$\mathbb{P}\left[|(s-x)S_t - s| \leq a\right] = \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)\right]$$

Substituting into the expression for the derivatives we get

$$\begin{aligned}
\frac{d\left(\phi(+,a) - \phi(-,a)\right)}{da} &= \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s+x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s+x}\right)\right] \\
&\quad - \left(\mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)\right]\right) \\
&= \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s+x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right)\right] \\
&\quad - \left(\mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s+x}\right)\right] - \mathbb{P}\left[W_t \leq \frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)\right]\right) \\
&= -\mathbb{P}\left[W_t \in I_1\right] + \mathbb{P}\left[W_t \in I_2\right],
\end{aligned}$$

where I_1(resp. I_2) is an interval of length $\ln\left(\frac{s+x}{s-x}\right)$ with endpoint $\frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right)$ (resp. $\frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)$). But because W_t is normally distributed this can only happen in two cases

(1) either $I_1 = I_2$ i.e. $\frac{1}{2}t + \ln\left(\frac{s+a}{s-x}\right) = \frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)$ implying $a = 0$,

(2) or I_1 and I_2 are symmetric, i.e. $\frac{1}{2}t + \ln\left(\frac{s+a}{s+x}\right) = -\left(\frac{1}{2}t + \ln\left(\frac{s-a}{s-x}\right)\right)$. The latter yields that $\ln(s^2 - a^2) = -t + \ln(s^2 - x^2)$ which eventually leads to $a^2 = s^2\left(1 - e^{-t}\right) + e^{-t}x^2$.

In any case the derivative $\frac{d(\phi(+,a)-\phi(-,a))}{da}$ can only be zero in two points, namely $a=0$ and $a=\sqrt{s^2\left(1-e^{-t}\right)+e^{-t}x^2}$. This allows us to apply the elementary lemma 6.5 above to the function $f(a)=\phi(+,a)-\phi(-,a)$. By lemma 6.2 we indeed have that $f(0)\geq 0$. The proof of lemma 6.6 is now complete. □

Proof of Theorem 6.1. We now have all the necessary material to complete the proof of the theorem 6.1. We put $\psi_k(x)=\Psi_k(x,s)$. From lemma 6.3, it follows that ψ_k has all the properties of lemma 6.6. This will be used in the series of inequalities below. Let $x\geq 0$ for simplicity of notation. The measure $\mathbb{Q}$ is defined as $d\mathbb{Q}=S_t\,d\mathbb{P}$.

$$\begin{aligned}
&\mathbf{E}\left[\Psi_k\left(x+s(S_t-1),sS_t\right)\right]\\
&\quad=\mathbf{E}\left[S_t\psi_k\left(xS_t^{-1}+s-sS_t^{-1}\right)\right]\\
&\quad=\mathbf{E}_{\mathbb{Q}}\left[\psi_k\left(x+(x-s)(S_t^{-1}-1)\right)\right]\\
&\quad=\mathbf{E}\left[\psi_k\left(x+(x-s)(S_t-1)\right)\right]\ \text{because of equality in law}\\
&\quad\leq\mathbf{E}\left[\psi_k\left(x+(x+s)(S_t-1)\right)\right]\ \text{because of lemma 6.6}\\
&\quad\leq\mathbf{E}_{\mathbb{Q}}\left[\psi_k\left(x+(x+s)(S_t^{-1}-1)\right)\right]\ \text{because of equality in law}\\
&\quad\leq\mathbf{E}\left[S_t\psi_k\left(x+(x+s)(S_t^{-1}-1)\right)\right]\\
&\quad\leq\mathbf{E}\left[\Psi_k\left(x-s(S_t-1),sS_t\right)\right].
\end{aligned}$$

The dynamic programming principle or a simple reasoning by induction now completes the proof. □

7. The Continuous Time Optimal Problem

In this section we will show how to derive the price of the passport option. The price will be obtained as a limit over discrete time optimization problems. We will also discuss the relation with optimal control theory. Throughout this section the process S will be defined on $[0,t_0]\times\Omega$, where $t_0<\infty$. As in section 6, we suppose that S is a geometric Brownian motion, i.e. $S_t=\exp\left(B_t-\frac{1}{2}t\right)$, where B is a Brownian motion with respect to a filtration $(\mathcal{F}_t)_{0\leq t\leq t_0}$. The price problem consists in calculating for given $x\in\mathbb{R}$,

$$\sup\left\{\mathbf{E}\left[\left(x+\int_{[0,t_0]}q_t\,dS_t\right)^+\right]\mid |q|\leq 1\text{ and predictable}\right\}.$$

As observed in section 2, this is equivalent to calculating the quantity

$$\sup\left\{\mathbf{E}\left[\left|x+\int_{[0,t_0]}q_t\,dS_t\right|\right]\mid |q|\leq 1\text{ and predictable}\right\}.$$

We first will give a solution in the case $x=0$. Since the piecewise constant strategies, i.e. the predictable integrands q, are dense, we can reduce the problem to a discrete time problem amd then pass to the limit. More precisely we introduce the sets, defined for $n\in\mathbb{N}$:

$$\mathcal{P}_n=\left\{q\mid |q|\leq 1, q\text{ constant on the intervals }\left]\frac{kt_0}{2^n},\frac{(k+1)t_0}{2^n}\right]\text{ and predictable}\right\}.$$

It is easily seen that $\cup_n \mathcal{P}_n$ is dense in the set $\mathcal{P}_\infty = \{q \mid |q| \le 1 \text{ and predictable}\}$, for the topology of convergence in measure with respect to $dm \otimes d\mathbb{P}$, where m is Lebesgue measure. By the dominated convergence theorem for stochastic integrals, or simply because S is an $\mathcal{H}^1$ martingale, we deduce that

$$\sup\left\{\mathbf{E}\left[\left|\int_{[0,t_0]} q_t\, dS_t\right|\right] \mid q \in \mathcal{P}_n\right\}$$

converges to

$$\sup\left\{\mathbf{E}\left[\left|\int_{[0,t_0]} q_t\, dS_t\right|\right] \mid q \in \mathcal{P}_\infty\right\}.$$

For each $n \in \mathbb{N}$, the optimal strategy is described by theorem 6.1 of section 6. The convergence result of lemma 5.2 in section 5 then implies that

$$\sup\left\{\mathbf{E}\left[\left|\int_{[0,t_0]} q_t\, dS_t\right|\right] \mid q \in \mathcal{P}_\infty\right\} = \mathbf{E}\left[\sup_{t\le t_0} S_t - 1\right].$$

The distribution of $\sup_{t\le t_0} S_t$ is easily calculated from the distribution of the supremum of a Brownian motion with drift. See [RY] page 70, Exercise 3.14 as well as p.320, exercise 1.21, for details. The case of $x \neq 0$ is treated as follows. The optimal strategies in discrete time all start (according to theorem 6.1) with the strategy $q = -\operatorname{sign}(x + S_t - 1)$. The distribution of the hitting time of this process with zero is known since it is the hitting time of a Brownian motion with drift (see again [RY] page 70, Exercise 3.14 as well as p.320, exercise 1.21). Once arrived at the point 0, the above obtained expression then gives the optimal value. For details of these calculations we refer to the papers [AAB], [HLP], where closed form expressions are given. These expressions can be derived from our expressions using a straightforward calculation. We omit the details.

We will now discuss the existence of an optimal strategy, for simplicity, we again assume that $x = 0$. The problem of finding an optimal strategy is closely related to the problem mentioned in the proof of theorem 5.1. It turns out that except in special cases (where the filtration $\mathcal{F}$ is big enough) there is no optimal strategy.

For a given strategy, i.e. a predictable process q such that $|q| \le 1$, we look at the equation $dX_t = q_t\, dS_t = q_t S_t\, dB_t$, where S is a geometric Brownian motion. We want to optimise $\mathbf{E}\left[|(q\cdot S)_{t_0}|\right]$. By the measure change $d\mathbb{Q} = S_{t_0}\, d\mathbb{P}$ (already used in section 6) we write this as $\mathbf{E}_{\mathbb{Q}}\left[|Y_{t_0}|\right]$, where $Y = X/S$ and by Itô's lemma, it follows the differential equation:

$$dY_t = (-Y_t + q_t)\, dW_t,$$

where according to Girsanov's theorem W is a Brownian motion under $\mathbb{Q}$. Stochastic control theory then shows that the optimal solution is given by

$$q_t = -\operatorname{sign}(Y_t) = -\operatorname{sign}(X_t).$$

We therefore have a look at the equation

$$dY_t = -\operatorname{sign}(Y_t)(|Y_t| + 1)\, dW_t.$$

The above transformations are quite standard and are present in one way or another in most of the papers ([AAB], [HLP], [Na], ...). In [Na] the discretisation is done at the level of the stochastic differential equation of the process Y. There is a subtle difference between this procedure and our discretisation. It is not clear whether the results of section 6 can be derived from the ones in [Na].
We now substitute $Z_t = f(Y_t)$ where f is the function $f(y) = \operatorname{sign}(y)\log(|y|+1)$. the function f is continuously differentiable and its second derivative is still a locally integrable function, namely $f''(y) = \frac{\operatorname{sign}(y)}{(|y|+1)^2}$. We can therefore apply Itô's lemma. We obtain

$$dZ_t = -\operatorname{sign}(Y_t)\left(dW_t + \frac{1}{2}dt\right) = -\operatorname{sign}(Z_t)\left(dB_t - \frac{1}{2}dt\right).$$

A measure change $d\mathbb{K} = \exp\left(\frac{1}{2}B_{t_0} - \frac{t_0}{8}\right) d\mathbb{P}$ then leads us to the equation

$$dZ_t = -\operatorname{sign}(Z_t)\, dR_t,$$

where $R_t = B_t - \frac{t}{2}$ is a Brownian motion under the measure $\mathbb{K}$. This equation is nothing else (provided we replace R by $-R$) but Tanaka's equation, see [RY] p. 358, exercise 1.19. It is known that if $\mathcal{F}$ is the filtration $\mathcal{F}^R$, generated by R, or which is the same, by B, then this equation has no strong solution. As a corollary we obtain that in this case *there is no optimal strategy for the passport option.* However the equation has a unique weak solution. If the filtration is big enough to host the process Z, then we can give an explicit form to the optimal process X. Indeed by transforming back, taking into account that $\operatorname{sign}(X) = \operatorname{sign}(Y) = \operatorname{sign}(Z)$, we get that

$$X_t = Y_t S_t = \operatorname{sign}(X_t)\left(\exp\left(|Z_t|\right) - 1\right)\exp\left(B_t - \frac{1}{2}t\right).$$

Exactly as before we can rewrite this expression using the local time of Z, or what is the same, the one sided maximal function of R. This gives

$$|Z_t| = -R_t + \sup_{s\le t} R_s,$$

which leads to the expression

$$X_t = \operatorname{sign}(X_t)\left(\exp\left(\sup_{s\le t} R_s\right) - \exp\left(R_t\right)\right).$$

Taking absolute values gives us that

$$|X_t| = \exp\left(\sup_{s\le t} R_s\right) - \exp\left(R_t\right) = \sup_{s\le t} S_s - S_t.$$

The filtration generated by the process $|X|$ is the same as the filtration generated by the starting Brownian motion B. The filtration generated by X requires extra randomness, which is given by the sign of the excursions of Z. Let us recall that by taking expected values, we find again the expression

$$\mathbf{E}\left[|X_t|\right] = \mathbf{E}\left[\sup_{s\le t} S_s\right] - 1.$$

The reader could ask whether the sequence of discrete time optimal strategies converges to the optimal solution in the continuous time case. The answer is no. Indeed this convergence would imply that the optimal solution would be predictable for the original filtration and hence would lead to the existence of a strong solution of Tanaka's equation. Also taking weak* limit points of the sequence of discrete time optimal solutions is leading nowhere. Such a weak* limit point might even be zero and the stochastic integrals are certainly not continuous with respect to the weak* topology. Taking convex combinations would therefore bring nothing useful. That the weak*-limit process is zero is something we did not check, although when the value process is starting at zero there is good evidence that this is indeed the case. So we are faced with the problem that there is no optimal strategy, but that nevertheless we can calculate the value function of the optimisation problem through either the discrete time approximation or through the concept of weak solution. Such a situation is quite general and in the case of Markov processes, Nisio, [Ni], showed that, under some regularity conditions, the value function in the continuous time case is the limit of the value functions of the discrete time approximations.

8. An extension to the case where interest rate is nonzero.

As mentioned in section 2, one variant of the passport option leads to the optimisation problem

$$\sup\left\{\mathbf{E}\left[\left|\int_0^{t_0} q_s e^{-fs}\,dS_s\right|\right] \mid |q|\le 1 \text{ and predictable}\right\}.$$

This problem can be solved exactly as in the case where $f=0$. Indeed the reasoning in sections 5 and 6 can be copied without any problem. The application of local time and Skorohod's lemma then yields that the quantity we have to calculate is

$$\mathbf{E}\left[\sup_{t\le t_0}\int_0^t e^{-fs}\exp\left(B_s-\frac{1}{2}s\right)\,dB_s\right].$$

To facilitate the calculations we introduce the notations

$$\begin{aligned}\nu &= -f-\frac{1}{2}\\ G_\nu(t) &= \exp\left(B_t+\nu t\right)\\ N_t^{(\nu)} &= \int_0^t G_\nu(s)\,dB_s\\ \Sigma_t^{(\nu)} &= \sup_{s\le t} N_s^{(\nu)},\end{aligned}$$

as well as the family of hitting times, defined for $a\ge 0$:

$$H_a^{(\nu)} = \inf\left\{t \mid N_t^{(\nu)} > a\right\}.$$

Clearly we have that $\left\{H_a^{(\nu)}\le t\right\} = \left\{\Sigma_t^{(\nu)}\ge a\right\}$. Also

$$\sup\left\{\mathbf{E}\left[\left|\int_0^{t_0} q_s e^{-fs}\,dS_s\right|\right] \mid |q|\le 1 \text{ and predictable}\right\} = \mathbf{E}\left[\Sigma_{t_0}^{(\nu)}\right].$$

We were not able to give a closed expression of this integral, neither could we find a better description of the law of $\Sigma_t^{(\nu)}$. But using an auxiliary exponential time we obtain an expression of the Laplace transform. More precisely we will calculate for each $\lambda > 0$, the integral

$$\int_0^\infty \mathbf{E}\left[\Sigma_t^{(\nu)}\right] \lambda \exp(-\lambda t)\, dt = \mathbf{E}\left[\Sigma_{\sigma_\lambda}^{(\nu)}\right],$$

where σ_λ is a random variable, independent of $\mathcal{F}_\infty$ and exponentially distributed with parameter $\lambda > 0$. In order to define such variable we might have to enlarge the probability space Ω. This is of course a standard procedure. The case $\nu = -1/2$ is treated in section 7 and the law of $\Sigma_t^{(-1/2)}$ is given by the law of the supremum of a Brownian motion with drift $-1/2$. As said before, there is a closed form description of this law. The general case $\nu \neq -1/2$ is harder. The result is given by:

Theorem 8.1. *The Laplace transform (with respect to λ) of the tail $\mathbb{P}\left[\Sigma_t^{(\nu)} > a\right]$, is given by:*

$$\mathbf{E}\left[\exp\left(-\lambda H_a^{(\nu)}\right)\right] = \mathbb{P}\left[\Sigma_{\sigma_\lambda}^{(\nu)} > a\right] = \lambda \int_0^\infty dt\, e^{-\lambda t} \mathbb{P}\left[\Sigma_t^{(\nu)} > a\right]$$
$$= \frac{1}{(a+1)^{\mu-\nu}} \frac{M\left(\alpha; 2\mu+1; \frac{2\nu+1}{a+1}\right)}{M(\alpha; 2\mu+1; 2\nu+1)},$$

where

$$\mu = \sqrt{2\lambda + \nu^2}$$
$$\alpha = \mu - \nu > 0$$

and where

$$M(x; y; z) = \int_0^1 dt\, e^{zt}\, t^{x-1} (1-t)^{y-x-1},$$

is the confluent hypergeometric function with parameters $x > 0$ and $y > x$. If we put $n_t^{(\nu)} = \mathbf{E}\left[\Sigma_t^{(\nu)}\right]$, then we have for all $\lambda > 0$ (in case $\nu \leq -1/2$) and for $\lambda > \frac{2\nu+1}{2}$ (in case $\nu \geq -1/2$) that:

$$\mathbf{E}\left[\Sigma_{\sigma_\lambda}^{(\nu)}\right] = \mathbf{E}\left[n_{\sigma_\lambda}^{(\nu)}\right] = \int_0^\infty n_t^{(\nu)} \lambda e^{-\lambda t}\, dt = \frac{1}{\mu+\nu+1} \frac{M(\alpha-1; 2\mu+1; 2\nu+1)}{M(\alpha; 2\mu+1; 2\nu+1)},$$

The proof is given through a series of transformations. The basic idea is to see the process G_ν as a time transformed Bessel process (Lamperti's identity). The same transformation then transforms the process $N^{(\nu)}$ into a Brownian motion. This will eventually transform the problem into a problem of finding the distribution of the first time a Bessel process hits a square root boundary. The details follow the same line of reasoning as in the case of Asian options, see [GY], [GY92], [Yor92], [Yor92a]. For a discussion of the confluent hypergeometric functions and the Bessel functions we refer to [W] and [Leb]. Confluent hypergeometric functions play an

important role in martingale theory as can be seen from the papers see [Dav], [Nov71], [Nov71a] and [Shepp], where the authors determine the best constants in inequalities relating the p^{th} moment of a stopped Brownian Motion W_T with the $(p/2)^{th}$ moment of the stopping time T.
Before starting the proof of Theorem 8.1, let us first give some remarks and relations with known results. We will restrict our attention to the case $\nu < 0$, which is for us the most relevant.

Corollary 8.2. *Let $\theta = -2\nu > 0$. Then one has*

$$\mathbb{P}\left[\Sigma_\infty^{(\nu)} > a\right] = \frac{1}{(a+1)^\theta} \frac{\int_0^1 dt\, t^{\theta-1} \exp\left(\frac{(2\nu+1)t}{a+1}\right)}{\int_0^1 dt\, t^{\theta-1} \exp\left((2\nu+1)\,t\right)} = \frac{\int_0^{\frac{1}{a+1}} dt\, t^{\theta-1} \exp\left((2\nu+1)\,t\right)}{\int_0^1 dt\, t^{\theta-1} \exp\left((2\nu+1)\,t\right)}.$$

Therefore

$$\Sigma_\infty^{(\nu)} \overset{\text{(law)}}{=} \frac{1}{U^{(\nu)}} - 1,$$

where $U^{(\nu)}$ is a random variable with density

$$\mathbb{P}\left[U^{(\nu)} \in dt\right] = \frac{dt\, t^{\theta-1} \exp\left((2\nu+1)\,t\right)}{\int_0^1 dt\, t^{\theta-1} \exp\left((2\nu+1)\,t\right)} \quad \textit{for } 0 < t < 1.$$

In the literature the distribution of $U^{(\nu)}$ is known as a truncated gamma distribution.

Proof. If in Theorem 8.1, we let λ tend to zero (which is the same as σ_λ tending to infinity), we find the required equality. The second equality is obtained by substituting $(2\nu + 1)t = u$ in the integral in the numerator. The equality in law is a restatement of this equality. □

We can also identify the distribution of $\Sigma_\infty^{(\nu)}$ in another way. This is the subject of

Corollary 8.3. *If $-1/2 \leq \nu < 0$, $\theta = -2\nu$, (and hence $0 < \theta \leq 2$) there is an identity in law:*

$$\Sigma_\infty^{(\nu)} \overset{\text{(law)}}{=} \frac{\mathbf{e}}{X^{(2\theta)}_{m_\nu,1/2}},$$

where $\mathbf{e}$ denotes a standard exponential distribution, independent of the denominator $X^{(2\theta)}_{m_\nu,1/2}$, which is the value of a BESQ process, with dimension 2θ, taken at time $1/2$ and starting from a truncated gamma distribution with density, defined for $0 < t < (2\nu+1)$, by $\mu_\nu(dt) = \frac{t^{\theta-1}e^t\,dt}{m_\nu}$, the number m_ν being a suitable normalisation constant. Consequently we have

$$U^{(\nu)} \overset{\text{(law)}}{=} \frac{X^{(2\theta)}_{m_\nu,1/2}}{\mathbf{e} + X^{(2\theta)}_{m_\nu,1/2}}.$$

Proof. If X is a BESQ process of dimension 2θ and starting at x, then the Laplace transform (with respect to a) of $X_{1/2}$ is given by

$$\mathbf{E}\left[\exp\left(-aX_{1/2}\right)\right] = (1+a)^{-\theta} \exp\left(\frac{-xa}{1+a}\right),$$

see the discussion after Corollary 1.3, Chap XI of [RY]. The result now follows by calculus.

Remark. If we take $\nu = -1/2$ in the equality of the above corollary 8.2, we find the well known fact that

$$\Sigma_\infty^{(-1/2)} \stackrel{\text{(law)}}{=} \frac{1}{U} - 1,$$

where U is a $[0,1]$-uniformly distributed random variable. The latter equality is easily proved using that $G^{(-1/2)}$ is a martingale, starting at 1 and tending to 0 at infinity. In corollary 8.3, we see that

$$\Sigma_\infty^{(-1/2)} \stackrel{\text{(law)}}{=} \frac{1}{U} - 1 \stackrel{\text{(law)}}{=} \frac{¢}{Y},$$

where Y is a BESQ process of dimension 2, starting at 0 (since $2\nu + 1 = 0$) and taken at time $1/2$. Since Y is a standard exponential, the result can also be verified directly.

Remark. If $0 > \nu \geq -1/2$ then the variable $\Sigma_\infty^{(\nu)}$ is no longer integrable. This means that the martingale $N^{(\nu)}$ is no longer in $\mathcal{H}^1$. This can be seen by calculating the quadratic variation of $N^{(\nu)}$. Indeed $\int_0^\infty \exp(2B_u + 2\nu u)\, du \geq \int_0^\infty \exp(2B_u - u)\, du$. The latter term is the quadratic variation when $\nu = -1/2$ and from Davis' theorem on $\mathcal{H}^1$ and the non integrability of $\Sigma_\infty^{(-1/2)}$ (the easy case), the statement follows. For $\nu \geq 0$ it is easily seen that the quadratic variation is ∞ and hence also $\Sigma_\infty^{(\nu)} = +\infty$.
The proof of Theorem 8.1 will be given in the next two sections.

8.1 On first hitting times of square root boundaries for Bessel processes.

The material in this section comes from [Yor84], see also [Shepp] for the case of Brownian motion instead of Bessel processes. For completeness, we give details. The reader not interested in the technicalities can skip the proofs.
Let R be a Bessel processes starting at the point ρ and of dimension $\delta = 2(\eta+1) \geq 0$. The law, on the space of all continuous functions, of this process is denoted by $\mathbb{P}_\rho^\eta$. For $c \geq \rho$, we also define two stopping times

$$\begin{aligned} T_c^+ &= \inf\left\{u \mid R_u = c\sqrt{1+u}\right\} \\ T_c^- &= \inf\left\{u \mid R_u = c\sqrt{1-u}\right\}. \end{aligned}$$

It is well known and easily seen that both stopping times are finite. Since the dimension $\delta \geq 2$ we also have that the Bessel process is defined up to time ∞. The distribution of the stopping times and more specifically from the hitting point is described by the following theorem:

Theorem 8.4. *We have the following expressions for the negative moments of the* *hitting points* $R_{T_c^{\pm}}$*:*

$$\mathbf{E}_\rho^\eta\left[\frac{1}{\left(R_{T_c^+}\right)^{2m}}\right]=\frac{1}{c^{2m}}\mathbf{E}_\rho^\eta\left[\frac{1}{\left(1+T_c^+\right)^m}\right]=\frac{1}{c^{2m}}\frac{M\left(m;\eta+1;\frac{\rho^2}{2}\right)}{M\left(m;\eta+1;\frac{c^2}{2}\right)}$$

$$\mathbf{E}_\rho^\eta\left[\frac{1}{\left(R_{T_c^-}\right)^{2m}}\right]=\frac{1}{c^{2m}}\mathbf{E}_\rho^\eta\left[\frac{1}{\left(1-T_c^-\right)^m}\right]=\frac{1}{c^{2m}}\frac{M\left(m;\eta+1;-\frac{\rho^2}{2}\right)}{M\left(m;\eta+1;-\frac{c^2}{2}\right)}$$

Proof of theorem 8.4. As explained in [Yor84], the main idea is to adapt Shepp's method, [Shepp], originally developed for the Brownian motion, to the case of Bessel processes. This is done using the fundamental martingales:

$$\tilde{I}_\eta\left(\theta R_u\right)\exp\left(-\frac{\theta^2 u}{2}\right)=\left(\frac{\theta R_u}{2}\right)^{-\eta}I_\eta\left(\theta R_u\right)\exp\left(-\frac{\theta^2 u}{2}\right)\text{ and}$$

$$\tilde{J}_\eta\left(\theta R_u\right)\exp\left(\frac{\theta^2 u}{2}\right)=\left(\frac{\theta R_u}{2}\right)^{-\eta}J_\eta\left(\theta R_u\right)\exp\left(\frac{\theta^2 u}{2}\right),$$

where the I_η and J_η are the standard Bessel functions, whence:

$$\tilde{I}_\eta(z)=\sum_{k=0}^{+\infty}\frac{(z/2)^{2k}}{\Gamma(k+1)\,\Gamma(k+\eta+1)}\text{ and}$$

$$\tilde{J}_\eta(z)=\sum_{k=0}^{+\infty}\frac{(-1)^k(z/2)^{2k}}{\Gamma(k+1)\,\Gamma(k+\eta+1)}$$

Since before times T_c^+ and T_c^- respectively, the martingales are bounded (see [Leb] e.g. for the necessary estimates on the behaviour of I and J), we may apply the optional sampling theorems at these times. We get that

$$\mathbf{E}_\rho^\eta\left[\tilde{I}_\eta\left(\theta c\sqrt{1+T_c^+}\right)\exp\left(-\frac{\theta^2}{2}T_c^+\right)\right]=\tilde{I}_\eta\left(\theta\rho\right)$$

$$\mathbf{E}_\rho^\eta\left[\tilde{J}_\eta\left(\theta c\sqrt{1-T_c^-}\right)\exp\left(\frac{\theta^2}{2}T_c^-\right)\right]=\tilde{J}_\eta\left(\theta\rho\right).$$

We now integrate both sides, on $\mathbb{R}_+$, with respect to the measure $\exp\left(-\frac{\theta^2}{2}\right)\theta^p\,d\theta$ and we get, after the obvious changes of variables

$$\theta'=\theta\sqrt{1+T_c^+}\text{ and }\theta''=\theta\sqrt{1-T_c^-},$$

the expressions

$$\mathbf{E}_\rho^\eta\left[\left(1+T_c^+\right)^{-\frac{1+p}{2}}\right]u_p(c)=u_p(\rho)$$

$$\mathbf{E}_\rho^\eta\left[\left(1-T_c^-\right)^{-\frac{1+p}{2}}\right]v_p(c)=v_p(\rho)$$

where

$$\int_0^\infty d\theta \, \exp\left(-\frac{\theta^2}{2}\right) \theta^p \tilde{I}(\theta c) = u_p(c)$$
$$\int_0^\infty d\theta \, \exp\left(-\frac{\theta^2}{2}\right) \theta^p \tilde{J}(\theta c) = v_p(c)$$

Looking up the representation of these functions, see [W], p 384–394 or [GR], formula 6.631 in the 5th edition, gives us that for $m = \frac{1+p}{2}$:

$$u_p(c) = \frac{\Gamma(m)}{\Gamma(\eta+1)} M\left(m; \eta+1; \frac{c^2}{2}\right) 2^{m-1}$$
$$v_p(c) = \frac{\Gamma(m)}{\Gamma(\eta+1)} M\left(m; \eta+1; \frac{-c^2}{2}\right) 2^{m-1}.$$

The proof of theorem 8.4 is therefore completed. □

8.2 The Proof of Theorem 8.1.

We now have the necessary ingredients to complete the proof of theorem 8.1. As stated in the beginning of section 8, we will time transform the process $N^{(\nu)}$ and $G^{(\nu)}$ into resp. a Brownian motion and a Bessel process. First, by Itô's rule, we write the stochastic differential equation for $G^{(\nu)}$:

$$G_t^{(\nu)} = 1 + N_t^{(\nu)} + \frac{2\nu+1}{2} \int_0^t ds \, G_s^{(\nu)} = 1 + \int_0^t G_s^{(\nu)} \, dB_s + \frac{2\nu+1}{2} \int_0^t ds \, G_s^{(\nu)}.$$

We therefore introduce the time transform function

$$A_t^{(\nu)} = \int_0^t ds \, \left(G_s^{(\nu)}\right)^2.$$

Clearly, we can now write $N^{(\nu)}$ as a time transformed Brownian motion, more precisely we have that

$$N_t^{(\nu)} = \gamma_{A_t^{(\nu)}},$$

where γ is a Brownian motion. At the same time, i.e. using the same time transform, we write

$$G_t^{(\nu)} = R^{(\nu)}_{A_t^{(\nu)}}$$

where $R^{(\nu)}$ is a Bessel process of dimension $\delta = 2(1+\nu)$, starting at 1 and given by the equation

$$dR_u = d\gamma_u + \frac{2\nu+1}{2} \frac{du}{R_u}$$

The inverse time transformation $C_u^{(\nu)} = \inf\left\{t \mid A_t^{(\nu)} \geq u\right\}$ will also play its role. It is clear that we may write

$$C_u^{(\nu)} = \int_0^u \frac{ds}{\left(R_s^{(\nu)}\right)^2}.$$

In case the index $\nu < 0$, the dimension of the Bessel process $R^{(\nu)}$ is stricly smaller than 2 and hence the Bessel process hits 0. So let us define $T_0^{(\nu)} = \inf\{t \mid R_t^{(\nu)} = 0\}$. For $\nu < 0$, we then have that $T_0^{(\nu)} < \infty$, almost surely. But this does not pose problems since $A_t^{(\nu)} < T_0^{(\nu)}$ for all $t < \infty$. More precisely we have that $A_\infty^{(\nu)} = T_0^{(\nu)}$, a fact that led the second author to a proof of Dufresne's identification of the distribution of $A_\infty^{(\nu)}$:

$$A_\infty^{(\nu)} \overset{\text{law}}{=} \frac{1}{2\gamma_\nu},$$

where γ_ν denotes a gamma distributed random variable, [Dufr],[Yor92], [Yor92a]. Also we can see that $C^{(\nu)}_{T_0^{(\nu)}} = \infty$.

For $\nu \geq 0$, we do not have such problems. In the sequel, the reader can check that all expressions we need, are not influenced by the fact that the Bessel processes can hit 0. Roughly speaking, we are only using the part of the trajectories before time $T_0^{(\nu)}$.

Step 1. *We have the following equality*

$$\mathbb{P}\left[\Sigma^{(\nu)}_{\sigma_\lambda} > a\right] = \mathbf{E}\left[\exp\left(-\lambda C^{(\nu)}_{T_a^*}\right)\right],$$

where $T_a^* = \inf\{u \mid \gamma_u > a\}$.

By definition we have that

$$\Sigma^{(\nu)} = \gamma^*_{A_t^{(\nu)}} \text{ where } \gamma_u^* = \sup_{s \leq u} \gamma_s$$

We therefore get that

$$\left\{\Sigma_t^{(\nu)} > a\right\} = \left\{\gamma^*_{A_t^{(\nu)}} > a\right\} = \left\{C^{(\nu)}_{T_a^*} \leq t\right\} \text{ and } C^{(\nu)}_{T_a^*} = H_a.$$

We now get

$$\begin{aligned}\mathbb{P}\left[\Sigma^{(\nu)}_{\sigma_\lambda} > a\right] &= \int_0^\infty dt\,\lambda e^{-\lambda t}\mathbb{P}\left[\Sigma_t^{(\nu)} > a\right] \\ &= \int_0^\infty dt\,\lambda e^{-\lambda t}\mathbb{P}\left[C^{(\nu)}_{T_a^*} \leq t\right] = \mathbf{E}\left[\exp\left(-\lambda C^{(\nu)}_{T_a^*}\right)\right].\end{aligned}$$

Step 2. *If, for each η, $\mathbb{P}_r^\eta$ denotes the law of the Bessel process $R_r^{(\eta)}$ of index η and starting at r, we have, for any stopping time of the "coordinate" process R:*

$$\mathbf{E}_r^{(\nu)}\left[\exp\left(-\lambda C_T\right)\right] = \mathbf{E}_r^{(\mu)}\left[\left(\frac{r}{R_T}\right)^{\mu-\nu}\right],$$

where $\mu = \sqrt{2\lambda + \nu^2} > 0$.

The proof of this equality can be found (at least for $\nu \geq 0$) in [Yor], page 77, formula 6.20. We include a sketch just for completeness, this proof is also valid in the case $\nu \leq 0$. We replace the stopping time T by the stopping times

$$T_N = \min\left(T, \inf\left\{u \mid \frac{R}{r} \leq \frac{1}{N}\right\}\right).$$

Afterwards we take the limit for $N \to \infty$. For the stopping times T_N we have to show that

$$\mathbf{E}_r^{(\nu)}\left[\exp\left(-\int_0^{T_N}\frac{du}{R_u^2}\right)\right] = \mathbf{E}_r^{(\mu)}\left[\left(\frac{r}{R_{T_N}}\right)^{\mu-\nu}\right].$$

This is an immediate consequence of Girsanov's theorem. We introduce a new probability measure, say $\mathbb{Q}$, defined as

$$d\mathbb{Q} = \exp\left(\int_0^{T_N}\frac{\alpha}{R_u}\,d\gamma_u - 1/2\int_0^{T_N}\left(\frac{\alpha}{R_u}\right)^2 du\right)d\mathbb{P}.$$

Here $\alpha = \sqrt{2\lambda+\nu^2}-\nu$, a constant that also later will play a role. The reader can check, that because we stopped the Bessel process before it hit the level $1/N$, this indeed defines a new probability measure. We can then write

$$\begin{aligned}&\mathbf{E}_r^{(\nu)}\left[\exp\left(-\int_0^{T_N}\frac{du}{R_u^2}\right)\right]\\ &\qquad = \mathbf{E}_{\mathbb{Q}}\left[\exp\left(-\int_0^{T_N}\frac{du}{R_u^2}\right)\exp\left(-\int_0^{T_N}\frac{\alpha}{R_u}\,d\gamma_u + 1/2\left(\frac{\alpha}{R_u}\right)^2 du\right)\right].\end{aligned}$$

Under the measure $\mathbb{Q}$, the process γ is turned into a Brownian motion with drift, i.e. $d\gamma_u = d\gamma'_u + \frac{\alpha\,du}{R_u}$, with γ' a $\mathbb{Q}$ Brownian motion (all processes stopped at T_N). Using Itô's lemma, we can then rewrite the right hand side as:

$$\mathbf{E}_r^{(\mu)}\left[\left(\frac{r}{R_{T_N}}\right)^{\mu-\nu}\right].$$

If we now take the limit for $N \to \infty$, we find on the left hand side

$$\mathbf{E}_r^{(\nu)}\left[\exp\left(-\lambda C_T\right)\right].$$

On the right hand side we may interchange the limit and the integral sign. To see this, look at the process $\left(\frac{r}{R_u}\right)^{2\mu}$ which is a local martingale under $\mathbb{P}_r^{(\mu)}$ (see [RY] p. 426, discussion about the speed measure). This martingale tends to zero at infinity (since $\mu > 0$). This yields an estimate on the maximum function of $\frac{r}{R_u}$. Using this we see that the sequence $\left(\frac{r}{R_{T_N}}\right)^{\mu-\nu}$ is uniformly integrable. The details are left to the reader.

Remark. There is a second choice for the parameter α that gives a similar relation. If we put $\alpha = -\mu-\nu$, then the same proof, together with a, justified, change of limit and integral, yields the following result:

$$\mathbf{E}_r^{(\nu)}\left[\exp\left(-\int_0^{T}\frac{du}{R_u^2}\right)\right] = \mathbf{E}_r^{(-\mu)}\left[\left(\frac{R_T}{r}\right)^{\mu+\nu}\right].$$

Remark that now the right hand side is an integral for a Bessel process that starts at r, but is of negative index (for big λ even of negative dimension). Such a process necessary hits zero, since $-\mu < 0$. Since the exponent is positive, this does not do any harm. The case $\mu = 0$ is of no interest since then necessarily $\lambda = \nu = 0$.

Step 3. *Reduction to a hitting time for a Bessel process*

The preceding two steps allowed us to reduce the proof of theorem 8.1 to a calculation of negative moments of the variable $R^{(\mu)}_{T^{(\mu)}_{\alpha,\beta}}$. Here R is a Bessel process of index μ starting at r and the stopping time $T^{(\mu)}_{\alpha,\beta}$ is defined as

$$T^{(\mu)}_{\alpha,\beta} = \left\{ u \mid R^{(\mu)} > \alpha + \beta \int_0^u \frac{ds}{R^{(\mu)}_s} \right\}.$$

The case of interest is where the starting point and the different parameters are given by $r = 1$, $\alpha = a + 1$, $\beta = c_\nu = \frac{2\nu+1}{2}$. The parameter β can be positive or negative and we will have to distinguish these two cases.

Step 4. *Reduction to a square root boundary.*

The idea is to time transform the process in such a way that the integral $\int_0^u \frac{ds}{R^{(\mu)}_s}$ becomes the new time. This is done using the following lemma, which is taken from [RY], Chap XI, proposition (1.11).

Lemma. *There is a Bessel process $\hat{R}^{(2\mu)}$, defined on the same probability space, of index 2μ starting at $2\sqrt{r}$, such that*

$$\left(R^{(\mu)}_u\right)^{1/2} = \frac{1}{2}\hat{R}^{(2\mu)}_{\int_0^u \frac{ds}{R^{(\mu)}_s}}.$$

Consequently we get that

$$R^{(\mu)}_{T^{(\mu)}_{\alpha,\beta}} = \frac{1}{4}\left(\hat{R}^{(2\mu)}_{\hat{T}_{\alpha,\beta}}\right)^2,$$

where the stopping time $\hat{T}_{\alpha,\beta}$ is defined as

$$\hat{T}_{\alpha,\beta} = \inf\left\{ u \mid \hat{R}^{(2\mu)} = 2\sqrt{\alpha + \beta u}\right\}.$$

Step 5. *Putting together the equalities.*

In the previous step we reduced the problem to a hitting time of a Bessel process of index 2μ with a square root boundary. We will now compare the stopping time of step 4 with the ones introduced in step 2. In order to do this we put $R'_u = \sqrt{\frac{|\beta|}{\alpha}}\hat{R}_{\frac{\alpha}{|\beta|}u}$. By the scaling property of Bessel processes (see [RY], property 1.10, page 427), the process R' is still a Bessel process of index 2μ, but starting at the point $\frac{r|\beta|}{\alpha}$. If we put $c = 2\sqrt{|\beta|}$, then it follows that

$$\hat{R}_{\hat{T}_{\alpha,\beta}} = \sqrt{\alpha}|\beta|R'_{T^+_c} \text{ if } \beta > 0$$
$$\hat{R}_{\hat{T}_{\alpha,\beta}} = \sqrt{\alpha}|\beta|R'_{T^-_c} \text{ if } \beta < 0.$$

This finally yields that

$$R^{(\mu)}_{T_{\alpha,\beta}} = \frac{\alpha}{4|\beta|}\left(\tilde{R}^{(2\mu)}_{T^+_{2\sqrt{\beta|}}}\right)^2 \text{ where } \tilde{R}^{(2\mu)}_0 = 2\sqrt{\frac{r|\beta|}{\alpha}}.$$

The results of step 2 now give us that, in case $\alpha \geq r$, $\mu > 0$ and for arbitrary β:

$$\mathbf{E}^{(\mu)}\left[\frac{1}{\left(R_{T_{\alpha,\beta}}\right)^m}\right] = \frac{1}{\alpha^m}\frac{M\left(m, 2\mu+1; 2r\frac{\beta}{\alpha}\right)}{M\left(m, 2\mu+1; 2\beta\right)}.$$

If we put $r = 1$, $\alpha = a + 1$ and $\beta = c_\nu = \frac{2\nu+1}{2}$, we get theorem 8.1.

Step 6. *The calculation of the Laplace transform of n_t^ν.*

This step is pure calculus. From the expression in theorem 8.1, we deduce that

$$\begin{aligned}\mathbf{E}\left[n^\nu_{\sigma_\lambda}\right] &= \int_0^\infty da\, \mathbf{E}\left[e^{-\lambda H_a^{(\nu)}}\right] \\ &= \int_0^\infty \frac{da}{(a+1)^\alpha} \frac{M\left(\alpha; 2\mu+1; \frac{2\nu+1}{a+1}\right)}{M\left(\alpha; 2\mu+1; 2\nu+1\right)}\end{aligned}$$

where $\alpha = \mu - \nu$. Hence

$$= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \int_0^1 dt\, t^{\alpha-1}(1-t)^{2\mu-\alpha} \int_0^\infty \frac{da}{(a+1)^\alpha} \exp\left(\frac{(2\nu+1)t}{a+1}\right)$$

which we write as

$$= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \int_0^1 dt\, t^{\alpha-1}(1-t)^{2\mu-\alpha} \int_1^\infty \frac{db}{b^\alpha} \exp\left(\frac{(2\nu+1)t}{b}\right)$$

or by putting $b = tu$ in the last integral

$$= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \int_0^1 dt\, (1-t)^{2\mu-\alpha} \int_{1/t}^\infty \frac{du}{u^\alpha} \exp\left(\frac{(2\nu+1)}{u}\right)$$

here we assumed that $\alpha > 1$ which is equivalent to $\lambda > \frac{2\nu+1}{2}$, a condition that is automatically satisfied if $\nu < -1/2$ and $\lambda > 0$. The expression can then be changed into

$$\begin{aligned}&= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \int_0^1 dt\, (1-t)^{\mu+\nu} \int_0^t \frac{dv}{v^{2-\alpha}} \exp\left((2\nu+1)v\right) \\ &= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \int_0^1 \frac{dv}{v^{2-\alpha}} \exp\left((2\nu+1)v\right) \int_v^1 dt\, (1-t)^{\mu+\nu} \\ &= \frac{1}{M\left(\alpha; 2\mu+1; 2\nu+1\right)} \frac{1}{\mu+\nu+1} \int_0^1 \frac{dv}{v^{2-\alpha}} \exp\left((2\nu+1)v\right)(1-v)^{\mu+\nu+1} \\ &= \frac{1}{\mu+\nu+1} \frac{M\left(\alpha-1; 2\mu+1; 2\nu+1\right)}{M\left(\alpha; 2\mu+1; 2\nu+1\right)}\end{aligned}$$

Remark. The reader can check that in the case $\nu = -1/2$, the expression of section 7 is found back. We omit the straightforward but unpleasant calculation.

References

[AAB] Andersen, L.; Andreasen, J. and Brotherton-Ratcliffe, R., *The Passport Option*, Aarhus University, Department of Mathematics, Working Paper (1998).

[Dav] Davis, B., *On the L^p norms of stochastic integrals and other martingales*, Duke Math. J. **43** (1976), 697–704.

[DM] Dellacherie, C. and Meyer, P.A., *Probabilités et potentiel, Chapitres V à VIII, Théorie des martingales*, Hermann, Paris (1980).

[DS94] Delbaen, F. and Schachermayer, W., *A General Version of the Fundamental Theorem of Asset Pricing*, Mathematische Annalen **300** (1994), 463–520.

[DS98] Delbaen, F and Schachermayer, W., *The Fundamental Theorem of asset pricing for unbounded stochastic processes*, Mathematische Annalen **312** (1998), 215–250.

[Dufr] Dufresne, D., *The distribution of a perpetuity, with applications to risk theory and pension funding*, Scand. Actuar. J. (1990), 39–79.

[GR] Gradshtein, I.S. and Ryzhik, I.M., *Table of Integrals, Series and Products, fifth Edition*, Academic Press (1994).

[GY] Geman, H. and Yor, M., *Quelques relations entre processus de Bessel, options asiatiques et fonctions confluentes hypergéométriques*, C.R.A.S. Paris, **T. 314, Série I** (1992), 471–474.

[GY92] Géman, H. and Yor, M., *Bessel processes, Asian Options and perpetuities*, working paper, Paris VI, Labo de probabilités (1992).

[HH] Henderson, V. and Hobson D.G., *Local Time, Coupling and teh Passport Option*, University of Bath, Working Paper (1998).

[HLP] Hyer T.; Lipton-Lifschitz A. and Pugachevsky D., *Passport to Success*, Risk Magazine **10** (1997), 127–131.

[Leb] Lebedev, N., *Special Functions and their Applications*, Dover (1972).

[Na] Nagayama, I., *Pricing of Passport Option*, J.Math.Sci.Univ.Tokyo **5** (1999), 747–785.

[Ni] Nisio, M., *Lectures on Stochastic Control Theory*, ISI Lecture Notes Series (1981).

[Nov71] Novikov, A.A., *On stopping times for a Wiener Process*, Theory of Proba. and its Appli. **16** (1971), 449–456.

[Nov71a] Novikov, A.A., *On moment inequalities for stochastic integrals*, Theory of Proba. and its Appli. **16** (1971), 538-541.

[Pr] Protter, Ph., *Stochastic Integration and Differential Equations, A new approach*, Springer (1990).

[RY] Revuz, D. and Yor, M., *Continuous Martingales and Brownian Motion, 2nd edition*, Springer (1994).

[Shepp] Shepp, L., *A first passage problem for the Wiener process*, Ann. Math. Statist. **38** (1967), 1912–1914.

[SV] Shreve, S. and Večeř, J., *Options on a Traded Account*, Finance and Stochastics (1999), forthcoming.

[W] Watson, G.N., *A Treatise on the Theory of Bessel Functions*, Cambridge University Press (1944).

[Yor] Yor, M., *Some Aspects of Brownian Motion*, Lectures in Mathematics ETH, Zürich, Birkhäuser (1992).

[Yor84] Yor, M., *On square root boundaries for Bessel processes and pole seeking Brownian motion*, Stochastic Analysis and Applications, Lecture Notes in Mathematics Nr. 1095, Springer (1984).

[Yor92] Yor, M., *Sur certaines fonctionnelles exponentielles du mouvement brownien réel*, J. Appl. Probab. **29** (1991), 202–208.

[Yor92a] Yor, M., *On some exponential functionals of Brownian Motion*, Adv. in Appl. Probab. **3**, 509–531.

Departement für Mathematik, Eidgenössische Technische Hochschule, ETH-Zürich, CH 8092 Zürich, Schweiz

Université Pierre et Marie Curie, Laboratoire de Probabilités, 4, Place Jussieu, Tour 56, F 75252 Paris Cedex 05, France

E-mail address: delbaen@math.ethz.ch