OPTIMAL STOPPING FOR SHEPP'S URN WITH RISK AVERSION

ROBERT CHEN[1], ILIE GRIGORESCU[1] AND MIN KANG[2]

ABSTRACT. An (m,p) urn contains m balls of value -1 and p balls of value $+1$. A player starts with fortune k and in each game draws a ball without replacement with the fortune increasing by one unit if the ball is positive and decreasing by one unit if the ball is negative. having to stop when $k=0$ (risk aversion). Let $V(m,p,k)$ be the expected value of the game. We are studying the question of the minimum k such that the net gain function of the game $V(m,p,k)-k$ is positive, in both the discrete and the continuous (Brownian bridge) settings. Monotonicity in various parameters m, p, k is established. Since the case $m-p<0$ is trivial, for $p\to\infty$, either $m-p\geq\alpha\sqrt{2p}$, when the the gain function cannot be positive, or $m-p<\alpha\sqrt{2p}$, when it is sufficient to have $k\sim\sqrt{p\log p}$, where α is a constant. We also determine an approximate optimal strategy with exponentially small probability of failure in terms of p. The problem goes back to Shepp [2], who determined the constant α in the unrestricted case when the net gain does not depend on k. A new proof of his result is given in the continuous setting.

1. INTRODUCTION

In [2], (also in [3], [4], and [6]) Shepp studied a series of optimal stopping problems. One of them is given in part a) of Section 6 of that paper, which contains a sketch of the proof, as follows.

An (m,p) urn contains m balls of value "-1" and p balls of value "$+1$", and the player is allowed to draw balls randomly without replacement until he wants to stop (he is also allowed not to draw at all). We are interested in finding the values for m and p such that

---

Date: March 15, 2014.

1991 *Mathematics Subject Classification.* Primary: 60G40; Secondary: 91A60, 93E20, 60C05.

Key words and phrases. Key words and phases: random urn, Chow and Robbins S_n/n problem, Brownian bridge, optimal strategy, optimal stopping.

[1] Department of Mathematics, University of Miami, 1365 Memorial Drive, Coral Gables, FL 33124-4250, email: igrigore@math.miami.edu.

[2] Department of Mathematics, North Carolina State University, SAS Hall, 2311 Stinson Dr., Box 8205, Raleigh, North Carolina 27695 email: kang@math.ncsu.edu.

there exists an optimal drawing policy for which the expected return (1.4), as a function of m and p, is positive. Let C denote the set of all (m,p) urns for which the expected return is positive.

Shepp reasons, by analogy with the main problem in his paper (the S_n/n problem of Chow and Robbins [1]) that there exists a sequence $\beta(1), \beta(2), ...$ of integers for which

$$C = \{(m,p) : m \leq \beta(p)\} . \tag{1.1}$$

Of course $\beta(p) \geq p$ and for large p we have

$$\lim_{p\to\infty} \frac{[\beta(p) - p]}{\sqrt{2p}} = \alpha \,, \tag{1.2}$$

where $\alpha = 0.83992...$ is the unique solution of the following integral equation explained later on in this paper in equation (3.22), more exactly

$$\alpha = (1 - \alpha^2) \int_0^\infty e^{\alpha t - t^2/2} dt.$$

The initial fortune k is not relevant for the construction of C in the *unrestricted model*, when the player would be able to play for any k, i.e. with unbounded debt. In this paper, we consider this urn scheme problem with *risk aversion*, which means that the player cannot go below a certain value of his fortune (cannot borrow money).

We introduce some notations. The process (k_n) of fortunes at times n, indexed by the discrete time units $n \geq 0$ is adapted to a filtration $(\mathcal{F}_n)_{n\geq 0}$. The assumption that the player cannot look into the future is formalized by requiring that any stopping strategy τ be a stopping time, i.e. $\{\tau > n\} \in \mathcal{F}_n$, $n \geq 0$. We notice that trivially $\tau \leq m + p$ and taking $z \in \mathbb{Z} \cup \{-\infty\}$ we define $\tau_0 = \inf\{n \geq 0 \,|\, k_n \leq z\}$ as the hitting time of the level set z, with the usual convention that the infimum over the empty set is $+\infty$.

We shall define $W(m, p, k; z)$ the value function of the game starting with fortune $k \geq 0$, $m \geq 0$ negative balls, $p \geq 0$ positive balls, and stopping when reaching fortune $z \geq -\infty$

$$W(m, p, k; z) = \sup_{\tau\in\mathcal{T}} E[k_\tau] \tag{1.3}$$

where $\mathcal{T}$ is the set of stopping times not exceeding the minimum of the hitting time of the level set $k = z$ and the final time $n = m+p$. In this case $\mathcal{T}$ is the *set of admissible strategies* of the game. In the discrete case, $z < -m$ is identical to $z = -\infty$. However, due to scaling, in the continuous case presented in Section 3 this setup is relevant.

Additionally $G(m, p, k; z) = W(m, p, k; z) - k$ is the expected return function of the game.

Inspecting the dynamic programming equation (2.1) we see that when $z = -\infty$ the return function is independent of k.

This is the case considered in [2] and then we may simplify the notation by setting

$$V(m,p) := G(m,p,k;-\infty)\,. \tag{1.4}$$

We retain the notation $V(m,p)$ in this case for consistency with the literature on the unrestricted model [2, 3, 4]. For example [3] adopts $k = 0$ since the initial fortune is not relevant, as mentioned.

In the risk aversion case, which is the subject of this paper, without loss of generality, we set $z = 0$. To simplify notation, since throughout Section 2 dedicated to the discrete case the value z is fixed, we denote

$$V(m,p,k) = W(m,p,k;0) \qquad \text{and} \qquad G(m,p,k) = G(m,p,k;0)\,. \tag{1.5}$$

We are interested in finding the smallest k for which the player makes a positive expected gain for a given (m,p) urn in C, i.e., finding the minimal k so that the expected value $V(m,p,k)$ exceeds k, when the player plays optimally. The answer can be summarized as follows:

(i) if $m < k$ then the unrestricted case and the risk-aversion case are not distinguishable and thus $G(m,p,k) > 0$ (Theorems 5 and 7);

(ii) if $m < p$, then $G(m,p,k) > 0$ for all $k \geq 1$;

(iii) if $m \geq p$, p is large and $m \geq p + \alpha\sqrt{2p}$, then $G(m,p,k) = 0$ for all k by (1.2) and Theorem 7;

(iv) if $m \geq k$, p is large and $p \leq m < p+\alpha\sqrt{2p}$, then a sufficient condition for $G(m,p,k) > 0$ is given in Theorem 1.

Case (iv) is the most difficult since it exhibits critical behavior, pointing naturally to the parabolic scaling (3.7) from the invariance principle for random walks.

The next result follows as a consequence of Theorem 12, which is, in its turn, based on the bounds established in Proposition 3.

Theorem 1. *Let $(m,p) \in C$ defined in (1.1). If p is large and $p \leq m < p + \alpha\sqrt{2p}$, then there exists $k(m,p) > 0$ such that $G(m,p,k) > 0$ if $k \geq k(m,p)$ and $G(m,p,k) = 0$ if*

$k < k(m,p)$. *An upper bound for* $k(m,p)$ *is given by* $c(p)\sqrt{2p}$, *with*

$$c(p) = \sqrt{c_1 + \frac{1}{2}\ln p}\,, \qquad c_1 = \frac{3}{2}\ln 2 - \frac{5}{2}\ln\alpha \approx 1.4758\,. \tag{1.6}$$

If, in addition, $m \leq p + \alpha_0\sqrt{2p}$, *for some* $\alpha_0 \in [0,\alpha)$, *then* $c(p)$ *has finite order, not exceeding the value* $x_*(\alpha_0)$ *defined in Proposition 3.*

The proof is at the end of the paper, after Theorem 13.

Remark. It is natural to look at $k(m,p)$ on a scale of order $\sqrt{m+p} \sim \sqrt{2p}$, since this corresponds to the scaling for the diffusive limit from Theorem 9. Asymptotically, only pairs (m,p) with $m \leq p + \alpha\sqrt{2p}$ belong to C. The value of $c = k(m,p)/\sqrt{2p}$ is bounded and depends only on $u = (m-p)/\sqrt{m+p}$ (see Proposition 3) provided that u remains bounded above away from α, say $u \leq \alpha_0 < \alpha$. It is only the *worst case scenario* $u \uparrow \alpha$ when $c \uparrow \infty$. In other words, given an arbitrary $(m,p) \in C$, the set of values allowing a positive gain in the unrestricted setting, (1.6) gives a value of k that ensures a positive gain for the risk-aversion setting. Numerical data shows that $c \leq 3$ for p from 2 to 1,000, which is consistent with the order of the approximation that predicts $c \approx 2.22$ with the logarithmic bound.

We give a brief outline of the paper. Section 2 gives monotonicity and comparison results between the unrestricted ($z = -\infty$) and the risk aversion ($z = 0$) case in the discrete setting via the recurrence corresponding to the dynamical programming equations for the optimal value (2.1). Section 3 treats the continuous limit of the problem, where under the diffusive scaling of Brownian motion (invariance principle) the underlying process is a Brownian bridge. Theorem 10 proves the relevant monotonicity results analogue to those of Section 2. Theorem 11 gives a new proof in the case $z = -\infty$ outlined by Shepp [2], while Theorem 12 gives an approximate winning strategy in the risk aversion case. In the discrete setting, Theorems 1 and 13 provide the asymptotic upper bounds for critical value of k for which a positive return is possible, as well as the approximate strategy to achieve it, via the continuous scaling limit.

2. Monotonicity results in the discrete case

The function $V(m,p,k)$ satisfies the following recursive relation,

$$V(m,p,k) = \max\left\{k, \frac{m}{m+p}V(m-1,p,k-1) + \frac{p}{m+p}V(m,p-1,k+1)\right\} \tag{2.1}$$

for all $m, p, k \geq 1$ with the following boundary conditions: For any $m, p \geq 0$, $V(m, p, 0) = 0$, for any $m, k \geq 0$, $V(m, 0, k) = k$, and for any $p \geq 0$, $k \geq 1$, $V(0, p, k) = p + k$. The optimal drawing policy is as follows: The game ends if the player does not have fortune any more or the first time $V(i, j, n) = n$ for some $n \geq 1$.

In [3] and [2], Boyce and Shepp proved the following three inequalities: $V(m+1, p) \leq V(m, p)$, $V(m, p) \leq V(m, p+1)$, and $V(m, p) \leq V(m+1, p+1)$ for all $m, p \geq 0$. Below we show that the function $V(m, p, k)$ exhibits similar monotonicity.

Theorem 2. *For any fixed $m, p \geq 0$, $V(m, p, k)$ is increasing in k.*

Proof. We will prove this inequality by induction on $t = m + p$. If either m or p equals 0, it is easy to see that $V(m, p, k)$ is increasing in k. From now on we will assume $m, p \geq 1$. Suppose $t = m + p = 2$, then $m = p = 1$. Then we have $V(1, 1, 0) = 0$, $V(1, 1, 1) = 1$ and $V(1, 1, k) = \frac{1}{2}(2k + 1)$ for all $k \geq 2$. So $V(m, p, k)$ is increasing in k when $t = m + p = 2$. Suppose that $V(m, p, k)$ is increasing in k for all $2 \leq t = m + p \leq s - 1$ for some $s \geq 3$. Now for $t = m + p = s$, we have

$$V(m, s-m, k) = \max\{k, \frac{m}{s} V(m-1, s-m, k-1) + (1 - \frac{m}{s})V(m, s-m-1, k+1)\}$$

$$\leq \max\{k+1, \frac{m}{s} V(m-1, s-m, k) + (1 - \frac{m}{s})V(m, s-m-1, k+2)\} = V(m, s-m, k+1).$$

Hence Theorem 1 also holds for $t = m + p = s$. Therefore, by mathematical induction, $V(m, p, k)$ is increasing in k for all $m, p \geq 0$. □

Theorem 3. *For any fixed $m, p, k \geq 1$, $V(m-1, p, k-1) \leq V(m, p-1, k+1)$.*

Proof. When $k = 1$, $V(m-1, p, k-1) = V(m-1, p, 0) = 0 < 2 \leq V(m, p-1, 2) = V(m, p-1, k+1)$. From now on, we will assume that $k \geq 2$. As in the proof of Theorem 2, we will use mathematical induction on $t = m + p$. First note that $t \geq 2$, since $m, p \geq 1$. Suppose that $t = 2$, then $m = 1$ and $p = 1$. $V(0, 1, k-1) = k < k + 1 = V(1, 0, k+1)$. Let $t = 3$. If $m = 2$ and $p = 1$, then $V(1, 1, k-1) = k - \frac{1}{2} \leq k + 1 = V(2, 0, k+1)$ for $k \geq 3$. For $k = 2$, we have $V(1, 1, k-1) = V(1, 1, 1) = 1 \leq 2 = k + 1 = V(2, 0, k+1)$. If $m = 1$ and $p = 2$, then $V(0, 2, k-1) = k + 1 \leq k + \frac{3}{2} = V(1, 1, k+1)$. So Theorem 3 holds for $t = 2, 3$.

Suppose that Theorem 3 holds for all $m, p \geq 1$ such that $t = m + p \leq s - 1$ where $s \geq 4$. Now let $t = m + p = s$. If $m = 1$, then $V(m-1, p, k-1) = V(0, p, k-1) = k + p - 1 \leq$

$k+p-1+\frac{1}{p} = V(1,p-1,k+1) = V(m,p-1,k+1)$, hence Theorem 3 holds. If $p=1$, then $V(m-1,p,k-1) = V(m-1,1,k-1) = k-1 \leq k+1 = V(m,0,k+1) = V(m,p-1,k+1)$, since $m \geq 3$ in this case. So Theorem 3 holds. From now on we will assume that $m,p \geq 2$ and $t = m+p = s \geq 4$.

For all $2 \leq m \leq s-2$, note that

$$V(m-1,s-m,k-1) =$$
$$\max\{k-1\,,\ \frac{m-1}{s-1}V(m-2,s-m,k-2) + \frac{s-m}{s-1}V(m-1,s-m-1,k)\},$$
$$V(m,s-m-1,k+1) =$$
$$\max\{k+1\,,\ \frac{m}{s-1}V(m-1,s-m-1,k) + \frac{s-m-1}{s-1}V(m,s-m-2,k+2)\}.$$

By induction hypothesis, we know $V(m-2,s-m,k-2) \leq V(m-1,s-m-1,k) \leq V(m,s-m-2,k+2)$, hence $V(m-1,s-m,k-1) \leq V(m,s-m-1,k+1)$. Now Theorem 2 also holds for $t = m+p = s$. Induction completes the proof. □

Theorem 4. *For all $m,p,k \geq 0$, $V(m,p,k) \leq V(m,p+1,k)$ and $V(m,p,k) \leq V(m-1,p,k)$.*

Proof. If one of m,p,k is 0, it is easy to check that Theorem 4 holds. So we will assume that $m,p,k \geq 1$. We will use induction argument on $t = m+p$. Suppose $t=2$, then $m=p=1$ since $m,p \geq 1$. If $k=1$, then $V(1,1,1) = 1 \leq 2 = V(0,1,1)$ and $V(1,1,1) = 1 \leq \frac{5}{3} = V(1,2,1)$. If $k \geq 2$, then $V(1,1,k) = k+\frac{1}{2} \leq k+1 = V(0,1,k)$ and $V(1,1,k) = k+\frac{1}{2} \leq k+\frac{4}{3} = V(1,2,k)$. Hence Theorem 4 holds for $t=2$. Assume that Theorem 4 holds for $t \leq s$ and assume that $t = m+p = s+1$, where $s \geq 2$. Note that

$$V(m,s+1-m,k) =$$
$$\max\left\{k, \frac{m}{s+1}V(m-1,s+1-m,k-1) + \frac{s+1-m}{s+1}\,V(m,s-m,k+1)\right\},$$

$$V(m-1,s+1-m,k) =$$
$$\max\left\{k, \frac{m-1}{s}V(m-2,s+1-m,k-1) + \frac{s+1-m}{s}\,V(m-1,s-m,k+1)\right\} \text{ and}$$

$$V(m,s+2-m,k) =$$
$$\max\left\{k, \frac{m}{s+2}V(m-1,s+2-m,k-1) + \frac{s+2-m}{s+2}\,V(m,s+1-m,k+1)\right\}.$$

Hence to compare $V(m-1,s+1-m,k)$ with $V(m,s+1-m,k)$, it is enough to look at the following difference,

$$\Big[\frac{m-1}{s}V(m-2,s+1-m,k-1)+\frac{s+1-m}{s}\,V(m-1,s-m,k+1)\Big]$$

$$-\Big[\frac{m}{s+1}V(m-1,s+1-m,k-1)+\frac{s+1-m}{s+1}\,V(m,s-m,k+1)\Big]$$

$$=\frac{m}{s+1}\,\big[V(m-2,s+1-m,k-1)-V(m-1,s+1-m,k-1)\big]$$

$$+\;\frac{s+1-m}{s+1}\,\big[V(m-1,s-m,k+1)-V(m,s-m,k+1)\big]$$

$$+\;\frac{s+1-m}{s(s+1)}\,\big[V(m-1,s-m,k+1)-V(m-2,s+1-m,k-1)\big]\;\geq\;0\,,$$

The reason we get this difference non-negative is the induction hypothesis that grants $V(m-2,s+1-m,k-1)-V(m-1,s+1-m,k-1)\geq 0$ and $V(m-1,s-m,k+1)-V(m,s-m,k+1)\geq 0$ and Theorem 3 implying $V(m-1,s-m,k+1)-V(m-2,s+1-m,k-1)\geq 0$. Hence $V(m,s+1-m,k)\leq V(m-1,s+1-m,k)$. The proof for $V(m,s+2-m,k)\geq V(m,s+1-m,k)$ is completely analogous. Therefore by mathematical induction, Theorem 4 holds for all $m,p,k\geq 0$. □

Boyce (in [3]) and Shepp (in [2]) proved that $V(m,p)\leq V(m+1,p+1)$ for any (m,p) urn. However, the statement that $V(m,p,k)\leq V(m+1,p+1,k)$ is not true in general. The equation (2.1) enables actual numerical computation and we see that, for example, $V(8,7,5)\approx 5.463869<5.469114\approx V(7,6,5)$ and $V(6,7,3)\approx 4.301282<4.409091\approx V(5,6,3)$.

Theorem 5. *For all $m,p\geq 0$ and $k\geq m+1$, $V(m,p)=V(m,p,k)-k$.*

Proof. It is easy to check that $V(m,p)=V(m,p,k)-k$ for all $m+p=t=0,1,2$ and $k\geq m+1$. Assume that Theorem 5 holds for all $t=m+p=0,1,2,...,s-1$. Then by induction hypothesis,

$$V(m,s-m)=\max\left\{0\,,\;\frac{m}{s}\,[V(m-1,s-m)-1]+\frac{s-m}{s}\,[V(m,s-m-1)+1]\right\}$$

$$= \max\left\{0, \frac{m}{s}\,[V(m-1, s-m, k-1)-(k-1)-1]+\frac{s-m}{s}\,[V(m, s-m-1, k+1)-(k+1)+1]\right\}$$

$$= \max\left\{0, \frac{m}{s}V(m-1, s-m, k-1) + \frac{s-m}{s}V(m, s-m-1, k+1) - k\right\}$$

$$= V(m, s-m, k) - k\,.$$

Therefore, Theorem 5 holds for all $m + p = t = 0, 1, 2, ..., s$. By mathematical induction, the proof is complete. $\square$

Recall that $G(m, p, k) = V(m, p, k) - k$ denotes the expected net profit under an optimal drawing policy. Below we show that the expected net profit, $G(m, p, k)$ is non-decreasing in k.

Theorem 6. *For all $m, p \geq 0$, $G(m, p, k)$ is non-decreasing in k.*

Proof. By Theorem 5, for all $m, p \geq 0$, $G(m, p, k) = V(m, p)$ for all $k \geq m + 1$. We will assume $0 \leq k \leq m$. It is easy to check that Theorem 6 holds if one of m, p, k equals 0. We will assume that $m, p \geq 1$ and $1 \leq k \leq m$. Then we have

$$G(m, p, k) = \max\left\{0\,, \frac{m}{m+p}\,V(m-1, p, k-1) + \frac{p}{m+p}\,V(m, p-1, k+1) - k\right\}. \tag{2.2}$$

Let $t = m+p$ be the induction parameter. First consider the case $t = 2$, then $m = p = 1$ since we assume $m, p \geq 1$. Then

$$G(1, 1, k) = \max\left\{0\,, \frac{1}{2}\,[V(0, 1, k-1) + V(1, 0, k+1)] - k\right\} = 0\,.$$

Hence Theorem 6 holds for $t = 2$. Now let us assume that Theorem 6 holds for for all $m, p \geq 1$ and $t = m + p = 2, 3, ..., s - 1$, where $s \geq 3$. Now for $t = s$, the equation (2.2) gives

$$G(m, p, k) = \max\left\{0\,, \frac{m}{s}\,V(m-1, s-m, k-1) + \frac{s-m}{s}\,V(m, s-m-1, k+1) - k\right\}$$

$$= \max\left\{0, \frac{m}{s}[V(m-1, s-m, k-1)-(k-1)]+\frac{s-m}{s}[V(m, s-m-1, k+1)-(k+1)]+\frac{s-2m}{s}\right\}$$

$$= \max\left\{0, \frac{m}{s}\,G(m-1, s-m, k-1) + \frac{s-m}{s}\,G(m, s-m-1, k+1) + \frac{s-2m}{s}\right\}. \tag{2.3}$$

On the other hands,

$$G(m,p,k+1) = \max\left\{0, \frac{m}{s}V(m-1,s-m,k) + \frac{s-m}{s}V(m,s-m-1,k+2) - (k+1)\right\}$$

$$= \max\left\{0, \frac{m}{s}[V(m-1,s-m,k)-k] + \frac{s-m}{s}[V(m,s-m-1,k+2)-(k+2)] + \frac{s-2m}{s}\right\}$$

$$(2.4) \qquad = \max\left\{0, \frac{m}{s}G(m-1,s-m,k) + \frac{s-m}{s}G(m,s-m-1,k+2) + \frac{s-2m}{s}\right\}.$$

From the induction hypothesis, we know that $G(m-1,s-m,k-1) \leq G(m-1,s-m,k)$ and $G(m,s-m-1,k+1) \leq G(m,s-m-1,k+2)$. Hence comparing the equations (2.3) and (2.4), we conclude $G(m,s-m,k) \leq G(m,s-m,k+1)$ and Theorem 6 holds for $t = m+p = s$. □

Theorem 7. *For all $m,p,k \geq 0$, $V(m,p) \geq G(m,p,k)$.*

Proof. By Theorem 5, $V(m,p) \geq G(m,p,k)$ holds if $k \geq m+1$. Now let us assume that $k \leq m$. It is easy to verify that Theorem 7 holds if one of m,p,k is 0, we will focus on the case $m,p,k \geq 1$. Let $t = m+p$ be the induction parameter. First notice that Theorem 7 holds for $t = 2$, in other words, for $m = p = 1$. Suppose that $V(m,p) \geq G(m,p,k)$ holds for all m,p such that $2 \leq t = m+p \leq s-1$ where $s \geq 3$. Then for $t = m+p = s$, the induction hypothesis gives

$$V(m,s-m) = \max\left\{0, \frac{m}{s}[V(m-1,s-m)-1] + \frac{s-m}{s}[V(m,s-m-1)+1]\right\}$$

$$\geq \max\left\{0, \frac{m}{s}[V(m-1,s-m,k-1)-(k-1)] + \frac{s-m}{s}[V(m,s-m-1,k+1)-(k+1)] + \frac{s-2m}{s}\right\}$$

$$= \max\left\{0, \frac{m}{s}V(m-1,s-m,k-1) + \frac{s-m}{s}V(m,s-m-1,k+1) - k\right\}.$$

Hence $V(m,s-m) \leq V(m,s-m,k) - k$ and the mathematical induction completes the proof. □

Theorem 8. *For all $m,p,k \geq 0$, $G(m,p,k) \leq G(m+1,p+1,k+1)$.*

Proof. It is easy to check if one of m, p, k is 0, Theorem 8 holds. We will assume that $m, p, k \geq 1$. For $t = m + p$ which will be used as the induction parameter, we first establish the inequality for all m, p such that $t = m + p \leq 3$ and for all $k \leq 3$. $G(1,1,1) = 0 < G(2,2,2) = \frac{1}{3}$ and $G(1,1,2) = \frac{1}{2} < G(2,2,3) = \frac{2}{3}$. $G(1,2,1) = \frac{2}{3} < G(2,3,2) = \frac{6}{5}$, $G(1,2,2) = \frac{4}{3} < G(2,3,3) = \frac{3}{2}$, $G(1,2,3) = \frac{4}{3} < G(2,3,4) = \frac{3}{2}$. $G(2,1,1) = G(3,2,2) = 0$, $G(2,1,2) = G(3,2,3) = 0$ and $G(2,1,3) = 0 < G(3,2,4) = \frac{1}{5}$.

Let us assume that Theorem 8 holds for all m, p such that $t \leq s - 1$ where $s \geq 4$. Now let $t = s$ and $k \leq s$, then

$$G(m, s-m, k) = \max\Big\{0, \frac{m}{s}G(m-1, s-m, k-1) + \frac{s-m}{s}G(m, s-m-1, k+1) + \frac{s-2m}{s}\Big\} \tag{2.5}$$

and on the other hand,

$$G(m+1, s-m+1, k+1) = \max\Big\{0, \frac{m+1}{s+2}G(m, s+1-m, k) + \frac{s+1-m}{s+2}G(m+1, s-m, k+2) + \frac{s-2m}{s+2}\Big\}. \tag{2.6}$$

First note that $G(m-1, s-m, k-1) \leq G(m, s+1-m, k)$ and $G(m, s-m-1, k+1) \leq G(m+1, s-m, k+2)$ from the induction hypothesis. Now we will show $G(m, s-m, k) \leq G(m+1, s-m+1, k+1)$ by dealing with the cases, $s < 2m$, $s = 2m$ and $s > 2m$ separately. If $s < 2m$, the difference of the second terms in the equations (2.6) and (2.5) is

$$\left[\frac{m+1}{s+2}G(m, s+1-m, k) + \frac{s+1-m}{s+2}G(m+1, s-m, k+2) + \frac{s-2m}{s+2}\right]$$

$$- \left[\frac{m}{s}G(m-1, s-m, k-1) + \frac{s-m}{s}G(m, s-m-1, k+1) + \frac{s-2m}{s}\right]$$

$$\geq \frac{s-2m}{s(s+2)}\big[G(m, s+1-m, k) - G(m+1, s-m, k+2) - 2\big] \geq 0. \tag{2.7}$$

To see the inequality in (2.7), using Theorem 3, note that

$$G(m+1, s-m, k+2) + 2 - G(m, s+1-m, k)$$

$$= V(m+1, s-m, k+2) - V(m, s+1-m, k) \geq 0.$$

If $s = 2m$, due to the induction assumptions, $G(m, m+1, k) \geq G(m-1, m, k-1)$ and $G(m+1, m, k+2) \geq G(m, m-1, k+1)$. Hence

$$G(m+1, s+1-m, k+1) - G(m, s-m, k) = G(m+1, m+1, k+1) - G(m, m, k)$$

$$= \frac{1}{2}\left[G(m, m+1, k) - G(m-1, m, k-1) + G(m+1, m, k+2) - G(m, m-1, k+1)\right] \geq 0\,.$$

Finally, if $s > 2m$, then note that

$$G(m+1, p+1, k+1) - G(m, p, k) = V(m+1, p+1, k+1) - V(m, p, k) - 1\,.$$

Here $p > m$ since $s = m + p > 2m$. By mathematical induction, we can prove that $V(m+1, p+1, k+1) - V(m, p, k) - 1 \geq 0$ if $m < p$. During the induction process, we will encounter the case that $p = m+1$. In that case

$$V(m+1, m+2, k+1) - V(m, m+1, k)$$

$$\begin{aligned} = \; & \frac{m+1}{2m+3} V(m, m+2, k) + \frac{m+2}{2m+3} V(m+1, m+1, k+2) \\ & - \frac{m}{2m+1} V(m-1, m+1, k-1) + \frac{m+1}{2m+1} V(m, m, k+1) \end{aligned}$$

By the induction hypothesis, $V(m, m+2, k) - V(m-1, m+1, k-1) \geq 1$. Similarly, since $V(m+1, m+1, k+2) = G(m+1, m+1, k+2) + k + 2$ and $V(m, m, k+1) = G(m, m, k+1) + k + 1$, $V(m, m+2, k) - V(m-1, m+1, k-1) \geq 1$ because $G(m+1, m+1, k+2) \geq G(m, m, k+1)$. Therefore, $G(m+1, p+1, k+1) - G(m, p, k) = V(m+1, p+1, k+1) - V(m, p, k) - 1 \geq 0$ if $s > 2m$. Hence the proof is complete by induction. $\square$

3. The continuous case

Let M_t be the number of balls marked with $-$ (minus), P_t the number of balls marked with $+$ (plus) and K_t the current amount of money (fortune) at time $t \geq 0$. We assume for the moment that the time is discrete.

A player picks a ball *without replacement*. If the ball is a minus, he loses one dollar, and if the ball is a plus, he wins one dollar. We specify that in this variant of the game, the player *must play each time*. This would not affect the answer to the question on the minimum initial fortune to guarantee a positive net profit $G(m, p, k)$. However, it simplifies the conservation laws (3.1).

The game can only last a maximum of $P_0 + M_0$ time units, but will stop as soon as $K_t = 0$. There exist constants A, B equal to the two conserved quantities in the problem

$$K_t - M_t + P_t = B\,, \qquad M_t + P_t + t = A\,. \tag{3.1}$$

The probability that K_t moves up by one unit is equal to $q_t = P_t/(P_t + M_t)$ and down by one unit to $1 - q_t$.

Assume the process (K_t) is defined on a probability space $(\Omega, \mathcal{F}, P)$ and is adapted to the filtration $(\mathcal{F}_t)$. Let τ_0 be the hitting time of the level set zero (running out of money).

We want to solve the optimal stopping problem

$$V(M_0, P_0, K_0) = \sup_{0 \le \tau \le A \wedge \tau_0} E[K_\tau]\,, \qquad \tau \quad \text{stopping time}\,. \tag{3.2}$$

A player can control the outcome by choosing the stopping time τ when to quit. The only restrictions are that he cannot play after running out of money and he cannot look into the future, which is formalized by having $\{\tau \le t\} \in \mathcal{F}_t$ for any $t \in [0, A]$.

The times of depletion of the pluses (respectively minuses) $\tau_\pm$ are such that $\tau_+ \wedge \tau_- \le A - 1$. For $\tau_+ \wedge \tau_- \le t \le A$ the fortune moves deterministically with $q_t = 0$, respectively $q_t = 1$.

This allows to write the relation

$$\begin{cases} Z_t = \frac{K_t - B}{A - t} = 1 - 2q_t = \frac{M_t - P_t}{P_t + M_t}\,, & \text{if} \quad 0 \le t < \tau_+ \wedge \tau_- \wedge \tau_0 \\ \qquad Z_t \in \{-1, +1\}\,, & \text{if} \quad \tau_+ \wedge \tau_- \wedge \tau_0 \le t \le \tau_0 \end{cases}, \tag{3.3}$$

where $+1$ is assumed when $\tau_+ < \tau_-$ and -1 in the opposite case (equality cannot happen). It follows that $q_t = \frac{1}{2}(1 - Z_t)$, leading to the inhomogeneous Markov chain

$$K_{t+1} = K_t \pm 1\,, \qquad \text{with probability} \quad \frac{1}{2}(1 \mp Z_t)\,. \tag{3.4}$$

Proposition 1. *The process (Z_t) is a martingale.*

Proof. Evidently $K_t \le K_0 + t < \infty$ and $A - t \ge 1$, τ_0 is a finite stopping time bounded above by a constant and the conditional expectations of the change in one step can be verified directly. □

It follows that

$$Z_0 = \frac{M_t - P_t}{P_t + M_t}\Big|_{t=0} = \frac{M_0 - P_0}{P_0 + M_0} = \frac{K_0 - B}{A}\,,$$

and since $K_t - B = (A-t)Z_t$, we can write

$$K_t = K_0(1-\frac{t}{A}) + B\frac{t}{A} + (A-t)(Z_t - Z_0)\,. \tag{3.5}$$

It is known [5] (6.24) that the Brownian bridge (X_t) pinned at (a,b), i.e. $X_a = b$ - almost surely and starting at $(0,k_0)$ satisfies

$$X_t = k_0(1-\frac{t}{a}) + b\frac{t}{a} + (a-t)\int_0^t \frac{dW_s}{a-s}\,. \tag{3.6}$$

One can see that Z_t is a discrete version of the martingale appearing in the stochastic integral. Let $N > 0$ and

$$k_t^N = N^{-1}K_{N^2 t}\,, \quad k_0 = N^{-1}K_0\,, \quad b = N^{-1}B\,, \quad a = N^{-2}A\,. \tag{3.7}$$

The parabolic scaling is natural to the Brownian motion. It can be regarded (considering $N \to \infty$) as an invariance principle for the original K_t with the understanding that $N^2 t^{macro} = t^{micro}$, $NK^{macro} = K^{micro}$. We define $(k_t^N)_{t\geq 0}$ at all lattice points $N^{-1}\mathbb{Z}$, $t = N^{-2}\mathbb{Z}_+$ and by linear interpolation between lattice points. The resulting paths are continuous. We then have the following invariance principle.

Theorem 9. *Under the scaling from (3.7), the process $(k_t^N)_{t\geq 0}$ converges weakly (in distribution) to the Brownian bridge, i.e. the process satisfying the SDE*

$$dX_t = \frac{b - X_t}{a-t}dt + dW_t\,, \qquad X_0 = k_0\,, \tag{3.8}$$

where $(W_t)_{t\geq 0}$ is a standard Brownian motion.

Proof. The proof is standard so we only sketch it. The scaling limit is proven by showing that $(k_t^N)_{t\geq 0}$, indexed by $N > 0$, is tight as a probability measure on $C([0, a-\delta], \mathbb{R})$. Here δ is arbitrary, strictly less than one. Any limit point is the law of a process solving the martingale problem for

$$\mathcal{L}f(t,x) = \partial_t f(t,x) + \frac{b-x}{a-t}f'(t,x) + \frac{1}{2}f''(t,x)$$

for all $f \in C_c^{1,2}([0, a-\delta]\times\mathbb{R}, \mathbb{R})$ (the subscript is for compact support in the space variable). It is a consequence of Doob's maximal inequality that the martingale part is tight. The drift term can be considered on any time interval away from a, i.e. $[0, a-\delta]$. Uniqueness of the solution to the martingale problem proves the result on $[0,a]$ by letting $\delta \downarrow 0$. $\square$

Denote the optimal function

$$w(a, b, k_0; z) = \sup_{0 \le \tau \le \tau_0 \wedge a} E[X_\tau]\,, \quad \tau_0 = \inf\{t > 0 | X_t \le z\} \tag{3.9}$$

for the Brownian bridge $(0, k_0) \to (a, b)$ stopped at the level set z.

Comparing to (3.1) we see that necessarily $M_t + P_t \sim O(N^2)$, $M_t - P_t \sim O(N)$. The scaling carries over to

$$p_t = N^{-2} P_t\,, \qquad m_t = N^{-2} M_t\,, \qquad m_t - p_t = N^{-1}(k_t - b) \tag{3.10}$$

via the equations $m_t - p_t = k_t - b$, $m_t + p_t = a - t$. The analogue problem to (3.2) corresponds to $z = 0$ and $t = 0$ and satisfies the scaling relation

$$v(m_0, p_0, k_0) := w(m_0 + p_0, k_0 - m_0 + p_0, k_0; 0) \sim N^{-1} V(M_0, P_0, K_0)\,. \tag{3.11}$$

Theorem 10. *The optimal function $v(m, p, k)$ and the gain function $g(m, p, k) = v(m, p, k) - k$ are (i) increasing in k when p and m are fixed; (ii) decreasing in $u = m - p$ when k and $m + p$ are fixed. (iii) When $z = -\infty$ the gain function $w(m + p, k - m + p, k; z) - k$ does not depend on k and is decreasing in $m - p$ when $m + p$ is fixed.*

Proof. The first observation is that $w(a, b, k; z)$ is increasing in b, $k = k_0$ and decreasing in z. In the variable z we simply have a larger set of stopping times. In the other variables, we use formula (3.6). The last term is independent of any parameters. The expected values are monotone for each stopping times τ in the optimization problem (3.9). It follows that the supremums are monotone in the respective parameter.

We notice that using the relation

$$w(a, b - d, k - d; z - d) + d = w(a, b, k; z)\,, \tag{3.12}$$

true for any real number d, for

(i) $d = k$ we obtain $v(m, p, k) = w(m + p, -(m - p), 0; -k) + k$ and $g(m, p, k) = w(m + p, -(m - p), 0; -k)$, which shows that v and g are increasing in k and (ii) decreasing in $u = m - p$ when k and $m + p$ are fixed;

(iii) $d = k - m + p$ we obtain $g(m, p, k) = w(m + p, 0, m - p; -k + (m - p)) - (m - p)$ which is increasing in k. When $z = -\infty$ the fourth component of the function is $-\infty$ and the gain function does not depend on k and is decreasing in $m - p$. □

The optimal function can be further simplified by noticing that

$$w(a,b,k_0;z) = a^{\frac{1}{2}} w(1, a^{-\frac{1}{2}}b, a^{-\frac{1}{2}}k_0; a^{-\frac{1}{2}}z) \tag{3.13}$$

and using (3.12). Putting $z = 0$ and $d = b$, $x = a^{-\frac{1}{2}}k_0$, $r = a^{-\frac{1}{2}}b$ and $u = x - r = a^{-\frac{1}{2}}(m_0 - p_0)$, we can see that

$$a^{\frac{1}{2}}(w(1,0,u;-r) + r) = w(a,b,k_0;0)\,, \tag{3.14}$$

which, in the original variables $a = p_0 + m_0 = p + m$, $b = b_0 = k_0 - m_0 + p_0 = k - m + p$, can be written as

$$v(m,p,k) = \sqrt{m+p}\; w\Big(1, \frac{k-m+p}{\sqrt{m+p}}, \frac{k}{\sqrt{m+p}}; 0\Big)\,. \tag{3.15}$$

In similar fashion, we have the gain function

$$g(m,p,k) = v(m,p,k) - k = a^{\frac{1}{2}}(w(1,0,u;-r) - u) \tag{3.16}$$

with the same notations.

An explicit formula (3.25) exists in [2] for the unrestricted resources case when the player does not face ruin, i.e. for $z = -\infty$. Let τ_c be the hitting time of the curve $c\sqrt{a-t}$ by the Brownian bridge X_t between the points $(0,u) \to (a,0)$, $a > 0$, $c \geq 0$ and

$$\bar{v}(a,u,c) = E^{(a,0)}_{(0,u)}[X_{\tau_c}\mathbf{1}_{\tau_c<a}]\,. \tag{3.17}$$

In the following we drop the initial and terminal point script when there is no danger of confusion.

It is proven in the same paper via (3.19) that when $c > ua^{-\frac{1}{2}}$, then $\tau_c < 1$ almost surely (evidently when $c = u$ we have $\tau_c = 0$). The Brownian bridge (3.8) admits a representation due to Doob of the form

$$X_t = b + \frac{B_s + k_0 - b}{1 + s/a}\,, \qquad s = \frac{t}{1 - t/a}\,, \quad 0 \leq t < a\,. \tag{3.18}$$

Based on the normalizations described in (3.13), we shall focus on the Brownian bridge X_t with $a = 1$. In terms of (3.18) with $a \to 1$, $k_0 \to u$ and $b \to 0$, the stopping time $s' = \tau_c/(1-\tau_c)$ will give the equation

$$B_{s'} + u = c\sqrt{1+s'}\,. \tag{3.19}$$

This permits an exact calculation, with the exact formula

$$\bar{v}(1,u,c) = cE[\sqrt{1-\tau_c}] = \frac{ch(1,u)}{h(1,c)} = \frac{c\,e^{\frac{u^2}{2}}\,\Phi(u)}{e^{\frac{c^2}{2}}\,\Phi(c)}\,, \tag{3.20}$$

where we introduced the notation

$$h(a,u) = \int_0^\infty e^{\lambda u - \frac{\lambda^2 a}{2}}\,d\lambda\,, \qquad -\infty < u < \infty\,, \quad a > 0\,. \tag{3.21}$$

Here $\Phi(z)$ denotes the distribution function of the standard normal. Then, directly by integration or using the Laplace transform of the complementary error function,

$$h(a,u) = a^{-\frac{1}{2}} h(1, a^{-\frac{1}{2}} u) \qquad \text{and} \qquad h(1,u) = \sqrt{2\pi} e^{\frac{u^2}{2}}\,\Phi(u)\,.$$

The function in (3.20) has a maximum in the variable c at $\alpha > 0$, equal to the solution of $h(1,\alpha) - \alpha h'(1,\alpha) = 0$, or equivalently

$$\alpha = (1-\alpha^2)\int_0^\infty e^{\lambda\alpha - \frac{\lambda^2}{2}}\,d\lambda\,. \tag{3.22}$$

Similarly $u \to (\alpha/h(1,\alpha))h(1,u) - u$ is well defined for all real u and convex. Its global minimum is zero and is achieved at $u = \alpha$. This is exactly the gain function in the case $z = -\infty$, i.e. $g(1,0,u) = w(1,0,u;-\infty) - u$ from (3.16) (using the same notation for simplicity). It has a double zero at $u = \alpha$

$$g(1,0,\alpha) = 0\,, \quad g'(1,0,\alpha) = 0\,, \quad g''(1,0,\alpha) = 2\alpha > 0\,, \tag{3.23}$$

where the last equality is obtained by direct computation.

Theorem 11 shows that, as long as $u \le \alpha$, the optimal function is the value of $\bar{v}$ at $c = \alpha$

$$w(1,0,u;-\infty) = \bar{v}(1,u,\alpha)\,. \tag{3.24}$$

More precisely, with the equality (3.22) in mind,

$$w(1,0,u;-\infty) = \begin{cases} (1-\alpha^2)\int_0^\infty e^{\lambda u - \frac{\lambda^2}{2}}\,d\lambda\,, & u \le \alpha \\ u & u > \alpha \end{cases}\,. \tag{3.25}$$

As discussed, the game allows an expected net gain $w(1,0,u;\infty) > u$ only when $u < \alpha$. In view of (3.14) we have

$$w(a,b,k;-\infty) = a^{\frac{1}{2}} w(1,0,a^{-\frac{1}{2}}(k-b);-\infty) + b \tag{3.26}$$

and again

$$(3.27) \qquad w(a,b,k;-\infty) = \frac{\alpha}{h(1,\alpha)} a h(a,k-b) + b\,.$$

The relation

$$(3.28) \qquad u \to w(1,0,u;-\infty) = \frac{\alpha h(1,u)}{h(1,\alpha)} \geq u\,, \qquad u \leq \alpha$$

shows that

$$(3.29) \qquad \frac{\alpha}{h(1,\alpha)} a^{\frac{1}{2}} h(1, a^{-\frac{1}{2}}(k-b)) + b = \frac{\alpha}{h(1,\alpha)} a h(a,k-b) + b \geq k\,.$$

Theorem 11. *The process*

$$t \to \frac{\alpha}{h(1,\alpha)} (a-t) h(a-t, k_t - b) + b\,, \qquad 0 \leq t \leq a$$

is a $(\mathcal{F}_t)$ - martingale. Due to (3.29), by comparing the values at $t = 0$ and $t = \tau$, a stopping time in the interval $[0,a]$, we obtain the upper bound

$$(3.30) \qquad \frac{\alpha}{h(1,\alpha)} a h(a,k-b) + b \geq w(a,b,k;-\infty)\,.$$

For initial $0 \leq a^{-\frac{1}{2}}(k-b) \leq \alpha$, the Brownian bridge $X_t = (k_t - b)/(a-t)$ reaches the curve $\alpha\sqrt{a-t}$ at $\tau_\alpha \in [0,a]$ with probability one and the bound is achieved for this stopping time, proving (3.30) holds with equality. For initial $a^{-\frac{1}{2}}(k-b) > \alpha$, the optimal function (3.26) is constant equal to k and is realized at $\tau = 0$.

Proof. We have to show that $t \to (a-t)h(a-t,k_t-b)$ is a martingale. The function is non-negative. Modulo a deterministic linear part in t, a Brownian Bridge has a mean zero Gaussian component represented by the stochastic integral in (3.6) with covariance $s \wedge t - sta^{-1}$. For a given $t \in [0,a]$ the variance is $t(a-t)a^{-1}$. Fubini's theorem applied to (3.21) when taking the expected value proves that $E[(a-t)h(a-t,k_t-b)]$ is always finite.

Ito's formula gives the time component

$$-h(a-t,k_t-b) - (a-t)\frac{\partial}{\partial a} h(a-t,k_t-b)$$

$$+(a-t)\Big[\frac{\partial}{\partial u} h(a-t,k_t-b)(-\frac{k_t-b}{a-t}) + \frac{1}{2}\frac{\partial^2}{\partial u^2} h(a-t,k_t-b)\Big]$$

$$= -\frac{\partial}{\partial u}[u h(a-t,u) - (a-t)\frac{\partial}{\partial u} h(a-t,u)] = 0\,, \qquad u = k_t - b\,,$$

after noticing that for any $a > 0$, $u \in \mathbb{R}$, the function h satisfies the equation

$$uh(a,u) - a\frac{\partial}{\partial u}h(a,u) = \int_0^\infty \frac{\partial}{\partial \lambda} e^{\lambda u - \frac{\lambda^2}{2}a} d\lambda = -1 \,.$$

From here we apply the optional stopping theorem. To prove that as soon as $a^{-\frac{1}{2}}(k - b) > \alpha$ we have $w(a,b,k;-\infty) = k$, we notice that in Theorem 10 (iii) the function $u \to w(1,0,u;-\infty) - u$ is non-increasing in u. Since at $u = \alpha$ the function vanishes, and the solution is trivially non-negative by adopting $\tau = 0$ in (3.9), we are done. $\square$

Let X_t be the Brownian bridge pinned at zero when $t = 1$, i.e. case $a = 1$ in (3.17). The expected value of the process at the stopping time $\tau_c \wedge \tau_0$ provides a lower bound for the optimal function $w(1,0,u;u-k)$.

Formula (3.41) p. 265 in [5] gives the distribution function of τ_0

$$F_{\tau_0}(s) = P_u(\tau_0 \le s) = 1 - \Phi(b\sqrt{t} + \frac{b+u}{\sqrt{t}}) + e^{-2(b+u)b}\Phi(b\sqrt{t} - \frac{b+u}{\sqrt{t}}) \,, \quad t = \frac{s}{1-s} \,. \tag{3.31}$$

Here and the next proposition we follow the convention from (3.17) that $E_u[\cdot] = E^{(1,b)}_{(0,u)}[\cdot]$.

Proposition 2. *Based on the distribution functions F_{τ_0} and F_{τ_c} we can calculate*

$$E_u[X_{\tau_c \wedge \tau_0} \mathbf{1}_{\tau_c < 1}] = \bar{v}(1,u,c) - \int_0^1 (\bar{v}(1-t,-b,c) + b) P_u(\tau_c > t) dF_{\tau_0}(t) \,, \tag{3.32}$$

with lower bounds $l_2 \ge l_1 \ge l_0$, in order of simplicity,

$$l_2 = \bar{v}(1,u,c) - \int_0^1 (c\sqrt{1-t} + b) P_u(\tau_c > t) dF_{\tau_0}(t) \,, \tag{3.33}$$

$$l_1 = \bar{v}(1,u,c) - (c+b)\int_0^1 P_u(\tau_c > t) dF_{\tau_0}(t) \,, \tag{3.34}$$

$$l_0 = \bar{v}(1,u,c) - (c+b)P_u(\tau_0 < 1) \,, \qquad P_u(\tau_0 < 1) = e^{-2(b+u)b} \,. \tag{3.35}$$

Remark. 1) The Laplace transform of F_{τ_c} can be calculated in the same way as (3.20) [2]. 2) Formula (3.20) calculates the expected value under the event that $\tau_c \le 1$ (the event $\tau_c = 1$ has zero probability), which was shown to have probability one when $0 \le u \le c$. However, the function $\bar{v}$ is well defined when $u < 0$. This is needed in the integrand of the second term of (3.32) where the starting point $-b$ after τ_0 may be negative.

Proof. Denote $\tau' = \tau_c \wedge \tau_0$ for simplification. In all relevant cases $c > 0$ and $\tau_c = \tau_0$ has zero probability. Then

$$E_u[X_{\tau'}\mathbf{1}_{\tau_c<1}] = E_u[X_{\tau'}\mathbf{1}_{\tau_c<1,\tau_c<\tau_0}] + E_u[X_{\tau'}\mathbf{1}_{\tau_c<1,\tau_c>\tau_0}] \tag{3.36}$$

$$= E_u[X_{\tau_c}\mathbf{1}_{\tau_c<1,\tau_c<\tau_0}] - bP_u(\tau_c < 1, \tau_c > \tau_0)$$

$$= \bar{v}(1,u,c) - E_u[X_{\tau_c}\mathbf{1}_{\tau_c<1,\tau_c>\tau_0}] - bP_u(\tau_c < 1, \tau_c > \tau_0)\,. \tag{3.37}$$

Since $c > u$ guarantees $\tau_c < 1$ with probability one, the probability in the last term subtracted equals

$$P_u(\tau_c < 1, \tau_c > \tau_0) = \int_0^1 P_u(\tau_c > t)dF_{\tau_0}(t) \le P_u(\tau_0 < 1) = e^{-2(b+u)b}\,, \tag{3.38}$$

where the upper bound was obtained from (3.31) at $t = +\infty$. Noting that $\{\tau_c > \tau'\} = \{\tau_c > \tau_0\}$, the other term subtracted in (3.37) equals

$$E_u[X_{\tau_c}\mathbf{1}_{\tau_c<1}\mathbf{1}_{\tau_c>\tau'}] = E_u[\mathbf{1}_{\tau_c>\tau'}E_u[X_{\tau_c\circ\theta_{\tau'}+\tau'}\mathbf{1}_{\tau_c\circ\theta_{\tau'}+\tau'<1}\,|\,\mathcal{F}_{\tau'}]]$$

$$= \int_0^1 \bar{v}(1-t,-b,c)P_u(\tau_c > t)dF_{\tau_0}(t)$$

which gives (3.32) after applying the strong Markov property for τ' since $\{\tau_c > \tau_c \wedge \tau_0\}$ is $\mathcal{F}_{\tau'}$ - measurable. □

We recall the gain function $g(m,p,k) = v(m,p,k) - k = w(a,b,k;0) - k = a^{\frac{1}{2}}(w(1,0,u;u-x) - u)$ from (3.16).

Proposition 3. *Let $\bar{g}(u,x) = (w(1,0,u;u-x) - u)_+$.*

(i) If $u \ge \alpha$, then $\bar{g}(u,x) = 0$ for any $x \ge 0$.

(ii) If $0 \le u < \alpha$, there exists $x_* = x_*(u) \ge u$ such that $\bar{g}(u,x) > 0$ for any $x > x_*$ and $\bar{g}(u,x) = 0$ for any $x \le x_*$. In this case, for any $c \ge u$, $x \ge u$

$$\bar{g}(u,x) \ge \bar{v}(1,u,c) - u - (c + x - u)e^{-2x(x-u)}\,, \tag{3.39}$$

and for any $c \in (u,\alpha]$, the lower bound from (3.39) seen as a function of x has a unique zero *$x_c \in (u,\infty)$, and the constant $x_* \le \inf_{c\in[u,\alpha]} x_c$. There exist values u for which the infimum over c is obtained at $c < \alpha$. Moreover, $\lim_{u\to\alpha^-} \inf_{c\in[u,\alpha]} x_c = +\infty$.*

Remark. 1) We may adopt $c = \alpha$ in (3.39) to obtain an estimate for x_*. The value obtained by (3.39) is not optimal.

2) The limit of the upper bound of $x^*(u)$ only says that our approximation (3.39) blows up as $u \uparrow \alpha$. It is a conjecture that, in fact, $\lim_{u \to \alpha^-} x^*(u) < +\infty$.

Proof. The stopping time $\tau' = \tau_c \wedge \tau_0$ provides a possible strategy and thus a lower bound for the gain function is $E_u[X_{\tau'}] - u$. This is provided by the exact calculation (3.36), noticing that $b = x - u$.

The integral in (3.32) is bounded above by

$$cP_u(\tau_0 < \tau_c) \leq cP_u(\tau_0 < 1) = ce^{-2(b+u)b} = ce^{-2x(x-u)}\,. \tag{3.40}$$

and the other term subtracted in (3.37) is exact.

Since $E_u[X_{\tau_c}] = E_u[X_{\tau_c}\mathbf{1}_{\tau_c<\tau_0}] + E_u[X_{\tau_c}\mathbf{1}_{\tau_c\geq\tau_0}]$ and $X_{\tau_c}\mathbf{1}_{\tau_c<\tau_0} = X_{\tau_c\wedge\tau_0}\mathbf{1}_{\tau_c<\tau_0}$, we have the inequality

$$E_u[X_{\tau_c\wedge\tau_0}\mathbf{1}_{\tau_c<\tau_0}] = E_u[X_{\tau_c}] - E_u[X_{\tau_c}\mathbf{1}_{\tau_c\geq\tau_0}] \geq$$

$$E_u[X_{\tau_c}] - cP_u(\tau_0 < 1) \geq \bar{v}(1, u, c) - cP_u(\tau_0 \leq 1)\,,$$

since $X_{\tau_c} = c\sqrt{1-\tau_c} \leq c$ and $\tau_c \leq 1$ a.s. On the other hand

$$E_u[X_{\tau_c\wedge\tau_0}] = E_u[X_{\tau_c\wedge\tau_0}\mathbf{1}_{\tau_c<\tau_0}] + E_u[X_{\tau_c\wedge\tau_0}\mathbf{1}_{\tau_c\geq\tau_0}] \geq$$

$$\bar{v}(1, u, c) - (c + (x-u)_+)P_u(\tau_0 \leq 1)\,.$$

Equation (3.40) p. 240 in [5] shows that $P_u(\tau_0 \leq 1) = e^{-2x(x-u)}$ as long as $x > 0$ and $x > u$. The function $x \to \bar{v}(1, u, c) - u - (c + x - u)e^{-2x(x-u)}$ in (3.39) is considered only for $x \geq u$. Since $\bar{v}(1, u, c) - c < 0$, it starts with a negative value at $x = u$. Its derivative $e^{-2x(x-u)}[(4x - 2u)(c + x - u) - 1]$ cannot have two zeros on $[u, \infty)$, otherwise the vertex of the parabola in brackets $(3u - 2c)/4 < u < c$ would also be greater than u. Because the derivative has at most one zero, the function starts with a negative value and ends as $x \to \infty$ with a positive value $\bar{v}(1, u, c) - u > 0$, it is analytic (the zeros are not dense), it follows that there exists exactly one zero in (u, ∞). Since $\bar{v}(1, u, x) - u > 0$ for $0 \leq u < \alpha$, we are done.

$\square$

It is easy to see that $u < 0$ (i.e. $m_0 < p_0$), then any strategy produces a win with probability one since the terminal point is zero, greater than u.

Let x_* be the constant obtained in (3.39) and c_* the value for which it is achieved.

Theorem 12. *A winning strategy for the continuum problem (3.9) with initial data* (m,p,k) *is to calculate* $u = \frac{m-p}{\sqrt{m+p}}$ *and*

(i) if $u \geq \alpha$ *then stop;*

(ii) if $0 \leq u < \alpha$*, then play until* $k_t - b = m_t - p_t \geq c_* \sqrt{m_t + p_t}$ *or* $k_t = 0$*, whichever occurs first, allowing the non-optimal but exponentially small probability of ruin.*

Proof. After time scaling with a factor of a, $(k_{at} - x)/\sqrt{a} \geq c\sqrt{1-t}$ is equivalent to $k_t - b = m_t - p_t \geq c\sqrt{m_t + p_t}$ for any c and thus for the optimum c_* as well.

It is clear from (3.32) that for any $c \geq u$

(3.41)

$$w(1,0,u;-b) - u \geq 0 \vee \bar{v}(1,0,u;-b) - u \,, \quad \text{where} \quad \bar{v}(1,0,u;-b) = \sup_{c\geq u} E_u[X_{\tau_c \wedge \tau_0}\, \mathbf{1}_{\tau_c \leq 1}] \,.$$

Using Proposition 3 we optimize over c obtaining c_*. □

Finally we can formulate an approximate winning strategy for the discrete problem (3.2), providing the best next move based on the current data.

Theorem 13. *Let* (M,P,K) *be the current data and assume* $M+P$ *is sufficiently large.*

Calculate $u = \frac{M-P}{\sqrt{M+P}} > 0$ *and*

(i) if $u \geq \alpha$ *then stop;*

(ii) if $0 \leq u < \alpha$*, then check if* $K > x_*(u)\sqrt{M+P}$*; if true continue, and if not, stop.*

Proof. The discrete case is an approximation based on the continuum case.

Let $\mathcal{X} \subseteq C([0,a],\mathbb{R})$ be the subset of continuous paths ω pinned at zero at time $t = a$ and $\mathcal{X}_0 \subseteq \Omega$ be the subspace of paths killed at the exit time $t_0 = t_0(\omega)$ from the open set bounded above by $x = c\sqrt{1-t}$ and below by $x = z$, i.e. $\omega(t) \equiv \omega(t_0)$ for all $t \geq t_0$. This time is finite, $t_0 \in [0,a]$ for $\omega \in \Omega$ since the upper boundary is itself in $\mathcal{X}$.

The trajectories of the discrete time process can be embedded in the space $\mathcal{X}_0$ by simple linear interpolation, in the standard construction used to prove Donsker's invariance principle. Under the scaling (3.7), the family of processes $\{k_t^N\}_{N>0}$ is tight (i.e. the laws are compact in the sense of Prokhorov's theorem). The limiting process (X_t), the Brownian bridge killed upon exiting the region at t_0, is also concentrated on the same path space $\mathcal{X}_0$. Now consider a functional Φ on $\mathcal{X}$ defined by $\Phi(\omega(\cdot)) := \omega(t_0) = \max\{c\sqrt{1-t_0}, z\}$. We note that Φ *is not continuous* on $\mathcal{X}$ and $\mathcal{X}_0$ is not a closed subspace of $\mathcal{X}$ in the supremum norm topology. However, convergence still holds. From Portmanteau's theorem on page 347 in [7], it follows that for a measurable functional $\Phi(\omega(\cdot))$ on Ω_0 we have

$$\lim_{N\to\infty} E[\Phi(k^N_\cdot)] = E[\Phi(X_\cdot)] \tag{3.42}$$

as soon as $\Phi(\cdot)$ is bounded and continuous at all points in $\mathcal{X}_0$ and the probability law P^X of the limiting process $(X_\cdot)$ satisfies $P^X(\mathcal{X}_0) = 1$.

It is sufficient to see that $\Phi(\omega(\cdot)) := \omega(t_0) = \max\{c\sqrt{1-t_0}, z\}$ is bounded by $c \vee |z|$ and is continuous on the set of killed paths $\mathcal{X}_0$. If two paths are within distance ϵ from each other in the uniform norm, based on the fact that the exit time from the region is actually in the interval $[0, a]$, we can pick the one exiting first. From there on, the other path must exit at a point which necessarily is within ϵ of the exit value of the first one, since the first path remained constant. The functional Φ defines the value of the game for both the discrete (k^N_t) and continuous limit (X_t). Since Proposition 3 proves a lower bound for the continuous limit, we obtained that the approximate strategy (i)-(ii) is a winning strategy as $N = M + P$ is large. □

We are ready to complete the proof of Theorem 1 in the discrete model version.

Proof of Theorem 1. 1) The existence of a critical $k(m, p)$ is a consequence of Theorem 6 which shows that the net gain function is nondecreasing in k. It is sufficient to find a value k with positive net gain. Part 2) of the proof will prove the existence of such a k and the upper bound for its minimal value as $p \to \infty$.

2) We denote $z = (m - p)/\sqrt{2p}$ and we have $z < \alpha$ in the hypothesis. The ratio $u = (m - p)/\sqrt{m + p}$ from Proposition 3 will then satisfy $u < \alpha$. This gives

$$u = u(z) = z\left(1 + \frac{z}{\sqrt{2p}}\right)^{-\frac{1}{2}} = z\left(1 - \frac{z}{2\sqrt{2p}} + \frac{3z^2}{16p} \ldots\right).$$

Let $x(u)$ be the solution of the equation in x of the lower bound from (3.39) set equal to zero. Any $c \geq \alpha$ in the formula would provide a lower bound, so we may take the extreme case $c = \alpha$. More precisely $x(u)$ is the solution of

$$\bar{v}(1, u, \alpha) - u - (\alpha + x - u)e^{-2x(x-u)} = 0\,, \qquad u = z\left(1 + \frac{z}{\sqrt{2p}}\right)^{-\frac{1}{2}}.$$

It is then sufficient to take

$$c(p) = \sup_{z<\alpha} x(u(z))\left(1 + \frac{z}{\sqrt{2p}}\right)^{\frac{1}{2}} \tag{3.43}$$

such that $k \geq c(p)\sqrt{2p}$ implies $G(m,p,k) > 0$. For large p, the function maximized in the formula for $c(p)$ is an increasing function in z, which leads to

$$c(p) = x(u(\alpha))(1 + \frac{\alpha}{\sqrt{2p}})^{\frac{1}{2}} ,$$

which is, as expected, the case when $m = p + \alpha\sqrt{2p}$. From (3.23) we know that $\bar{v}(1,u,\alpha) - u \approx \frac{1}{2}g''(1,0,\alpha)(\alpha - u)^2$, with $g''(1,0,\alpha) = 2\alpha$, which gives the asymptotic value of $x(u)$, now a function of p only,

$$\alpha\epsilon^2 - (\epsilon + x)e^{-2x(x-\alpha+\epsilon)} = 0 , \quad \epsilon = \frac{\alpha^2}{2\sqrt{2p}} .$$

The function in the equation is increasing in x, giving a sufficient lower bound $x \geq x_\delta$ where x_δ is the solution of the simpler equation $\alpha\epsilon^2 - e^{-(2-\delta)x^2} = 0$, for any $\delta > 0$ sufficiently small. Inverting x_δ, letting $\delta \downarrow 0$ and using the relation between ϵ and p we obtain the sufficient condition

$$c(p) \geq \sqrt{-\frac{\ln\alpha}{2} + \ln\frac{2\sqrt{2}}{\alpha^2} + \frac{1}{2}\ln p} ,$$

which implies the lower bound (1.6).

3) To prove the second bound of the theorem when $m \leq p + \alpha_0\sqrt{2p}$ we simply notice that in that case $c(p)$ defined in (3.43) has limit $x_*(\alpha_0)$ as $p \to \infty$ because the supremum is taken over values $z \leq \alpha_0 < \alpha$.

4. Open problems

The summary of the current results given above Theorem 1 points out to a a few open questions. Assume $m \geq p$ and p is large. If $m \geq p + \alpha\sqrt{2p}$, then $V(m,p) = 0$ and $V(m,p,k) - k = 0$. We concentrate on the case $p \leq m < p + \alpha\sqrt{2p}$ corresponding to (iv). In this case we know that $V(m,p) > 0$. Since

$$0 \leq V(m,p,k) - k \leq V(m,p) , \tag{4.1}$$

a question is: What are the largest values of k for which 1) the first inequality is an equality, and then 2) the second inequality is strict, respectively? Denote by $k_1(p)$ and $k_2(p)$ the two values. If $k \leq k_1(p)$ the second equality cannot take place, so $k \leq k_2(p)$, implying that $k_1(p) \leq k_2(p)$. We know that $k > m$ implies $k > k_2(p)$, so we have

$$k_1(p) \leq k_2(p) \leq m < p + \alpha\sqrt{2p} .$$

Theorem 2 proves that $k_1(p)$ is at most of order $\sqrt{p\ln p}$.

Suppose $\gamma \in (0, \alpha)$ and $m \approx p + \gamma\sqrt{2p}$. Then what are k_1, k_2 seen as functions of γ?

Some numerical computation seems to support that, in fact, $G(m, p, k) < V(m, p)$ if $k \leq m$, conjecturing that $k_2(p) = m \approx p + \alpha\sqrt{2p}$ whenever $(m, p) \in C$. However we don't have a proof for the strict inequality, $G(m, p, k) < V(m, p)$ if the (m, p) urn is in C.

Acknowledgements. We wish to thank to Burton Rosenberg for valuable discussions related to the computational aspects of the problem.

References

[1] Y. S. Chow and H. Robbins, *On optimal stopping rules for S_n/n*, Illinois J. Math., Vol. 9, pp. 444 - 454 (1965).

[2] L. A. Shepp, *Explicit Solutions to Some Problems of Optimal Stopping*, Ann. Math. Statist., Vol. 40, pp. 993 - 1010 (1969).

[3] W. M. Boyce, *On a Simple Optimal Stopping Problem*, Discrete Math., Vol. 5, pp. 297 - 312 (1973).

[4] R. W. Chen and F. K. Hwang, *On the Values of an (m,p) Urn*, Congressus Numerantium, Vol. 41, pp. 75 - 84 (1984).

[5] I. Karatzas and S. E. Shreve, *Brownian Motion and Stochastic Calculus*, Graduate Texts in Mathematics, Vol. 113. Springer-Verlag, New York (1991).

[6] R. W. Chen, A. M. Odlyzko, L. A. Shepp and A. Zame, *An Optimal Acceptance Policy for an Urn Scheme*, Siam J. Discrete Math., Vol. 11, pp. 183 - 195 (1998).

[7] P. Mörters and Y. Peres, *Brownian Motion*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, Cambridge (2010).