A Discrete-Time Homing Problem with Two Optimizers

Mario Lefebvre

Department of Mathematics and Industrial Engineering, Polytechnique Montréal, 2500, Chemin de Polytechnique, Montréal, QC H3T 1J4, Canada; mlefebvre@polymtl.ca

Abstract: A stochastic difference game is considered in which a player wants to minimize the time spent by a controlled one-dimensional symmetric random walk $\{X_n, n = 0, 1, \ldots\}$ in the continuation region $C := \{1, 2, \ldots\}$, and the second player seeks to maximize the survival time in C. The process starts at $X_0 = x > 0$ and the game ends the first time $X_n \leq 0$. An exact expression is derived for the value function, from which the optimal solution is obtained, and particular problems are solved explicitly.

Keywords: random walk; first-passage time; homing problem; difference game; dynamic programming; difference equation

1. Introduction

A deterministic one-dimensional two-player linear-quadratic (nonzero-sum) difference game can be defined as follows (see, for example, [1]):

$$X_{n+1} = a_n X_n + b_n u_n + c_n v_n \tag{1}$$

for $n \in \{0, 1, \ldots, T\}$, where a_n, b_n and c_n are deterministic functions, the control u_n (respectively, v_n) gives the decision of player 1 (respectively, player 2) at time n. Each player has a general quadratic cost function that he/she tries to minimize.

The difference game can be made stochastic by adding the random variable ϵ_n in Equation (1). The random variables $\epsilon_0, \epsilon_1, \ldots$ are assumed to be independent and identically distributed.

The final time T can be finite or infinite. Reddy and Zaccour [2] considered a class of non-cooperative N-player finite-horizon linear-quadratic dynamic games with linear constraints. Lin [3] studied the Stackelberg strategies in the infinite horizon LQ mean-field stochastic difference game. Liu et al. [4] used an adaptive dynamic programming approach to solve the infinite horizon linear quadratic Stackelberg game problem for unknown stochastic discrete-time systems with multiple decision makers. Ju et al. [5] used dynamic programming to obtain an optimal linear strategy profile for a class of two-player finite-horizon linear-quadratic difference games.

In this paper, we consider the one-dimensional controlled Markov chain $\{X_n, n = 0, 1, \ldots\}$ defined by

$$X_{n+1} = X_n + u_n - v_n + \epsilon_n \quad \text{for } n = 0, 1, \ldots, \tag{2}$$

where $\epsilon_n = \pm 1$ with probability $1/2$ and the random variables $\epsilon_0, \epsilon_1, \ldots$ are independent, so that $\{X_n, n = 0, 1, \ldots\}$ is a (controlled) symmetric random walk. The chain starts at $X_0 = x > 0$.

We define the first-passage time

$$T(x) := \inf\{n > 0 : X_n \leq 0 \mid X_0 = x > 0\}. \tag{3}$$

Our aim is to find the controls $u_n \in \{-1, 0\}$ and $v_n \in \{0, 1\}$ that minimize the expected value of the cost function

$$J(x) = \sum_{n=0}^{T(x)-1} (u_n^2 - v_n^2 + \lambda), \tag{4}$$

where λ is a positive constant. This parameter gives the penalty incurred for survival in the continuation region $C := \{1, 2, \ldots\}$. It is needed to obtain a well-defined problem. Indeed, if we set $\lambda = 0$, then the optimal solution for the first (respectively, second) player is trivially to choose $u_n \equiv 0$ (respectively, $v_n \equiv 1$).

Thus, there are two optimizers. The first one, using u_n, would like the Markov chain to hit the origin as soon as possible. Therefore, he/she would like to choose $u_n = -1$, but this generates a cost. On the other hand, the second optimizer wants the Markov chain to remain positive as long as possible. Hence, he/she would prefer to choose $v_n = 0$, which however generates no costs. Both optimizers, and especially the second one, must also take into account the value of the constant λ.

Remark 1. *We could, in theory, assume that the parameter λ is negative. Then, it is mainly the first optimizer who would need to consider the value of λ. See, however, Remark 3.*

The main difference between the current paper and the related ones found in the literature is the fact that, in our case, the final time is neither finite or infinite; it is rather a *random variable*.

The above problem is a particular *homing problem*, in which a stochastic process is controlled until a certain event occurs. This type of problem was introduced by Whittle ([6] p. 289) for n-dimensional diffusion processes. He also considered the case when we take the risk-sensitivity of the optimizer into account; see [7], as well as [8,9].

The author has written numerous papers on homing problems. In [10,11], these problems where extended to the case of discrete-time Markov chains, whereas in [12] the case of autoregressive processes was treated; see also [13].

In [10,11], the authors considered a problem related to the one defined above, but with only one optimizer. Thus, they treated a stochastic optimal control problem, whereas the problem in the current paper is a stochastic difference game.

To solve our problem, we will use dynamic programming. Let $F(x)$ be the value function defined by

$$F(x) = \min_{\substack{(u_n, v_n) \\ n = 0, \ldots, T(x) - 1}} E[J(x)]. \tag{5}$$

In the next section, the dynamic programming equation satisfied by the function $F(x)$ will be derived.

2. Dynamic Programming

In theory, we must determine the optimal value of (u_n, v_n) for $n = 0, 1, \ldots, T(x) - 1$. However, using dynamic programming, the problem is reduced to finding the optimal solution at the initial time $n = 0$ only.

Indeed, we can write, making use of Bellman's principle of optimality, that

$$
\begin{aligned}
F(x) \ &:=\ \min_{\substack{(u_n, v_n) \\ n=0,\ldots,T(x)-1}} E\left[\sum_{n=0}^{T(x)-1}(u_n^2 - v_n^2 + \lambda)\right] \\[2mm]
&=\ \min_{\substack{(u_n, v_n) \\ n=0,\ldots,T(x)-1}} \left\{u_0^2 - v_0^2 + \lambda + E\left[\sum_{n=1}^{T(x)-1}(u_n^2 - v_n^2 + \lambda)\right]\right\} \\[2mm]
&=\ \min_{(u_0, v_0)} \left\{u_0^2 - v_0^2 + \lambda + E[F(x + u_0 - v_0 + \epsilon_0)]\right\}.
\end{aligned}
\tag{6}
$$

We can now state the following proposition.

Proposition 1. *The value function $F(x)$ satisfies the dynamic programming equation*

$$
F(x) = \min_{(u_0, v_0)} \left\{u_0^2 - v_0^2 + \lambda + \frac{1}{2}[F(x + u_0 - v_0 - 1) + F(x + u_0 - v_0 + 1)]\right\}.
\tag{7}
$$

The equation is subject to the boundary condition

$$
F(x) = 0 \quad \text{if } x \le 0.
\tag{8}
$$

Remark 2. *The usefulness of the value function is that it enables us to determine the optimal controls $u^*(0)$ and $v^*(0)$.*

Since $(u_0, v_0) \in \{(-1, 0), (-1, 1), (0, 0), (0, 1)\}$, we deduce from Equation (7) that

$$
\begin{aligned}
F(x) \ =\ \min\Big\{&1 + \lambda + \frac{1}{2}[F(x - 2) + F(x)], \lambda + \frac{1}{2}[F(x - 3) + F(x - 1)], \\[2mm]
&\lambda + \frac{1}{2}[F(x - 1) + F(x + 1)], -1 + \lambda + \frac{1}{2}[F(x - 2) + F(x)]\Big\}.
\end{aligned}
\tag{9}
$$

Now, let $T_0(x)$ be the random variable that corresponds to $T(x)$ when $u_n = v_n \equiv 0$. Using the well-known results on the *gambler's ruin problem* (see, for instance, ([14] p. 349), we can state that

$$
E[T_0(x)] = \infty \quad \text{for any } x > 0.
\tag{10}
$$

Hence, we would also have

$$
E[J(x)] = \infty \quad \text{for any } x > 0.
\tag{11}
$$

Since the objective is to minimize the expected value of $J(x)$, we must conclude that the optimal solution is not $(u_0, v_0) = (0, 0)$.

Remark 3. *(i) We deduce from what precedes that we cannot choose a value of the parameter λ in the interval $(-\infty, 0)$, otherwise we obtain an infinite expected reward by choosing $u_n = v_n \equiv 0$.*

(ii) If $\lambda = 0$, then there is no penalty (or reward) for survival in the continuation region $\{1, 2, \ldots\}$. The optimal solution is obviously $u_n^* \equiv 0$. As will be seen below, the optimal value of v_n is 1 for any value of $\lambda \ge 0$.

(iii) If we define

$$
T(x) := \inf\{n > 0 : X_n \le 0 \text{ or } X_n \ge d \mid X_0 = x > 0\},
\tag{12}
$$

then, when $u_n = v_n \equiv 0$, we find that (see, again, ([14] p. 349))

$$E[T_0(x)] = x(d - x) \quad for\ 0 \le x \le d, \tag{13}$$

so that

$$E[J(x)] = \lambda\, x(d - x) \quad for\ 0 \le x \le d. \tag{14}$$

Therefore, we could take $\lambda < 0$ in that case.

(iv) If $P[\epsilon_n = 1] \equiv p$ and $P[\epsilon_n = -1] \equiv q$ and $T(x)$ is defined as in (12), then (see ([14] p. 348))

$$E[T_0(x)] = \frac{x}{q - p} - \frac{d}{q - p}\frac{1 - (q/p)^x}{1 - (q/p)^d} \quad for\ 0 \le x \le d \tag{15}$$

and

$$\lim_{d \to \infty} E[T_0(x)] = \begin{cases} \infty & if\ p \ge 1/2, \\ \frac{x}{q-p} & if\ p < 1/2. \end{cases} \tag{16}$$

Thus, we could consider the case when $\lambda < 0$ if $p < 1/2$.

Next, notice that $(u_0, v_0) = (-1, 0)$ and $(u_0, v_0) = (0, 1)$ yield the same expected value of X_1. However, the choice $(u_0, v_0) = (-1, 0)$ generates a *cost* of 1, while with $(u_0, v_0) = (0, -1)$ a *reward* of 1 is obtained. Hence, taking $(u_0, v_0) = (0, 1)$ is surely a better decision than choosing $(u_0, v_0) = (-1, 0)$. Thus, we must determine whether the optimal solution is $(u_0, v_0) = (-1, 1)$ or $(0, 1)$.

Proposition 2. *We deduce from what precedes that the second optimizer should choose $v_n \equiv 1$, independently of the value of u_n.*

Remark 4. *The optimal choice for u_n will depend on the value of the parameter λ, which, as mentioned above, gives the penalty incurred for survival in the continuation region.*

It follows from Proposition 2 that Equation (9) can be simplified to

$$F(x) = \min\left\{ \lambda + \frac{1}{2}[F(x - 3) + F(x - 1)], -1 + \lambda + \frac{1}{2}[F(x - 2) + F(x)] \right\}. \tag{17}$$

Proposition 3. *The value function $F(x)$ satisfies the non-linear third-order difference equation*

$$\begin{aligned}
& 2F^2(x) - F(x)\left[6\lambda - 4 + F(x - 1) + 2F(x - 2) + F(x - 3)\right] \\
& + F(x - 2)\left[F(x - 1) + F(x - 3) + 2\lambda\right] + 2(\lambda - 1)\left[F(x - 1) + F(x - 3)\right] \\
& + 4\lambda(\lambda - 1) = 0
\end{aligned} \tag{18}$$

for $x = 1, 2, \dots$ The boundary condition is

$$F(x) = 0 \quad if\ x \le 0. \tag{19}$$

Proof. Making use of the formula

$$\min\{a, b\} = \frac{a + b - |a - b|}{2}, \tag{20}$$

we deduce from Equation (17) that

$$\begin{aligned}
& 2F(x) - \left\{ -1 + 2\lambda + \frac{1}{2}[F(x) + F(x - 1) + F(x - 2) + F(x - 3)] \right\} \\
& = -\left| 1 + \frac{1}{2}[F(x - 1) + F(x - 3) - F(x) - F(x - 2)] \right|.
\end{aligned} \tag{21}$$

By squaring both sides of the above equation and simplifying, we obtain Equation (18). □

Solving a boundary value problem for a non-linear difference equation of order 3 is not easy. Instead of trying to solve Equation (18) directly, we will proceed as in [10].

3. Optimal Choice for u_n

We must determine whether the first optimizer should take $u_0 = -1$ or $u_0 = 0$. The control variable u_n (as well as v_n) is actually a function of X_n.

Suppose that we set $u_n(X_n) \equiv -1$. Then, denoting the function $F(x)$ by $\Phi(x)$, Equation (17) implies that

$$\Phi(x) = \lambda + \frac{1}{2}[\Phi(x-3) + \Phi(x-1)], \tag{22}$$

that we rewrite as follows:

$$2\Phi(y+3) - \Phi(y+2) - \Phi(y) = 2\lambda, \tag{23}$$

where $y := x - 3$. The equation is valid for $y \in \{-2, -1, \ldots\}$ and is subject to the boundary condition

$$\Phi(y) = 0 \quad \text{if } y \leq -3. \tag{24}$$

Making use of the mathematical software program *Maple*, we find that the solution of Equation (23) that satisfies the boundary conditions $\Phi(-3) = \Phi(-4) = \Phi(-5) = 0$ is

$$\Phi(y) = \frac{\lambda}{112}\left[210 + 56y + (i\sqrt{7} - 7)\left(\frac{i\sqrt{7} - 1}{4}\right)^y - (i\sqrt{7} + 7)\left(\frac{i\sqrt{7} - 1}{2}\right)^{-y}\right] \tag{25}$$

for $y \geq -3$, so that

$$\Phi(x) = \frac{\lambda}{112}\left[42 + 56x - (i\sqrt{7} + 21)\left(\frac{i\sqrt{7} - 1}{4}\right)^x + (i\sqrt{7} - 21)\left(\frac{i\sqrt{7} - 1}{2}\right)^{-x}\right] \tag{26}$$

for $x \geq 0$.

Remark 5. *(i) The function $\Phi(x)$ is real, even if it contains the imaginary constant i.*

(ii) Because Equation (23) is a third-order linear difference equation, we need three boundary conditions to solve it uniquely. Therefore, we used Equation (24) for $y = -3, -4$ and -5.

Next, let us denote the function $F(x)$ by $\Psi(x)$ if we set $u_n(X_n) \equiv 0$. We then deduce from Equation (17) that the function $\Psi(x)$ satisfies the second-order difference equation

$$\Psi(x) = -1 + \lambda + \frac{1}{2}[\Phi(x-2) + \Phi(x)]. \tag{27}$$

The unique solution that is such that $\Psi(0) = \Psi(-1) = 0$ is

$$\Psi(x) = (-1 + \lambda)\left[x - \frac{(-1)^x}{2} + \frac{1}{2}\right] \quad \text{for } x \geq 0. \tag{28}$$

Every time the value of X_n changes, the (first) optimizer must make a new decision. It then follows from Equation (17) that we can express the value function $F(x)$ in terms of $\Phi(x)$ and $\Psi(x)$ as in the following proposition.

Proposition 4. *The value function $F(x)$ is given by*

$$F(x) = \min\left\{\lambda + \frac{1}{2}[\min\{\Phi(x-3), \Psi(x-3)\} + \min\{\Phi(x-1), \Psi(x-1)\}],\right.$$
$$\left. -1 + \lambda + \frac{1}{2}[\min\{\Phi(x-2), \Psi(x-2)\} + \min\{\Phi(x), \Psi(x)\}]\right\} \quad (29)$$

for $x = \in \{0, 1, \ldots\}$.

To determine the value function, and hence the optimal value of $u_0 (= u_0(x))$, we can compare the two expressions

$$G(x) := \lambda + \frac{1}{2}[\min\{\Phi(x-3), \Psi(x-3)\} + \min\{\Phi(x-1), \Psi(x-1)\}] \quad (30)$$

and

$$H(x) := -1 + \lambda + \frac{1}{2}[\min\{\Phi(x-2), \Psi(x-2)\} + \min\{\Phi(x), \Psi(x)\}]. \quad (31)$$

That is, we can write that

$$F(x) = \min\{G(x), H(x)\}. \quad (32)$$

Remark 6. *We must set $\Psi(-2) = 0$ when we compute the function $G(x)$ (and $F(x)$).*

In the next section, we will present the results obtained with different values of the parameter λ to see the effect of this parameter on the optimal control $u_0^*(x)$.

4. Numerical Examples

Assume first that $\lambda = 1/2$. Table 1 gives the value function $F(x)$, $\Phi(x)$, $\Psi(x)$, $G(x)$, $H(x)$ and the optimal control $u_0^*(x)$ for $x = 1, 2, \ldots, 10$. We see that the optimal control is always $u_0^* = 0$, which could have been expected because the penalty for survival in the continuation region $C := \{1, 2, \ldots\}$ is not large enough to incite the first optimizer to use $u_0 = -1$ in order to leave C as rapidly as possible.

Table 1. Functions $F(x)$, $\Phi(x)$, $\Psi(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 1, 2, \ldots, 10$ when $\lambda = 1/2$.

x	$F(x)$	$\Phi(x)$	$\Psi(x)$	$G(x)$	$H(x)$	$u_0^*(x)$
1	−1	0.5	−1	0.5	−1	0
2	−1	0.75	−1	0	−1	0
3	−2	0.875	−2	0	−2	0
4	−2	1.1875	−2	−1	−2	0
5	−3	1.4687	−3	−1	−3	0
6	−3	1.6719	−3	−2	−3	0
7	−4	1.9297	−4	−2	−4	0
8	−4	2.1992	−4	−3	−4	0
9	−5	2.4355	−5	−3	−5	0
10	−5	2.6826	−5	−4	−5	0

Next, we take $\lambda = 1$. This case is rather special, because the function $\Psi(x)$ becomes $\Psi(x) \equiv 0$. Again, $u_0^*(x) \equiv 0$; see Table 2.

Table 2. Functions $F(x)$, $\Phi(x)$, $\Psi(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 1, 2, \ldots, 10$ when $\lambda = 1$.

x	$F(x)$	$\Phi(x)$	$\Psi(x)$	$G(x)$	$H(x)$	$u_0^*(x)$
1	0	1	0	1	0	0
2	0	1.5	0	1	0	0
3	0	1.75	0	1	0	0
4	0	2.375	0	1	0	0
5	0	2.937	0	1	0	0
6	0	3.344	0	1	0	0
7	0	3.859	0	1	0	0
8	0	4.398	0	1	0	0
9	0	4.871	0	1	0	0
10	0	5.365	0	1	0	0

In Table 3, $\lambda = 2$. This time, $u_0^*(x)$ is not a constant: $u_0^*(1) = -1$ or 0 and then $u_0^*(x) = -1$ for x odd and $u_0^*(x) = 0$ for x even.

Table 3. Functions $F(x)$, $\Phi(x)$, $\Psi(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 1, 2, \ldots, 10$ when $\lambda = 2$.

x	$F(x)$	$\Phi(x)$	$\Psi(x)$	$G(x)$	$H(x)$	$u_0^*(x)$
1	2	2	2	2	2	-1 or 0
2	2	3	2	3	2	0
3	3	3.5	4	3	3.75	-1
4	4	4.75	4	4.75	4	0
5	5	5.875	6	5	5.687	-1
6	6	6.687	6	6.687	6	0
7	7	7.719	8	7	7.797	-1
8	8	8.797	8	8.797	8	0
9	9	9.742	10	9	9.730	-1
10	10	10.730	10	10.730	10	0

Finally, $\lambda = 10$ in Table 4 and the optimal control is $u_0^*(x) \equiv -1$, which is not surprising because the constant λ is large.

Table 4. Functions $F(x)$, $\Phi(x)$, $\Psi(x)$, $G(x)$ and $H(x)$, and optimal control $u_0^*(x)$ for $x = 1, 2, \ldots, 10$ when $\lambda = 10$.

x	$F(x)$	$\Phi(x)$	$\Psi(x)$	$G(x)$	$H(x)$	$u_0^*(x)$
1	10	10	18	10	14	-1
2	15	15	18	15	16.5	-1
3	17.5	17.5	36	17.5	22.75	-1
4	23.75	23.75	36	23.75	28.375	-1
5	29.375	29.375	54	29.375	32.437	-1
6	33.437	33.437	54	33.437	37.594	-1
7	38.594	38.594	72	38.594	42.984	-1
8	43.984	43.984	72	43.984	47.711	-1
9	48.711	48.711	90	48.711	52.652	-1
10	53.652	53.652	90	53.652	57.818	-1

4.1. Critical Value of λ

For a fixed x, we can determine the critical value of λ, that is, the value of λ for which $G(x) = H(x)$. For instance, we see in Figure 1 that $G(1) = H(1)$ when $\lambda = 2$. For $\lambda \in (0, 2)$, $u_0^*(1) = 0$, while $u_0^*(1) = -1$ for $\lambda > 2$.
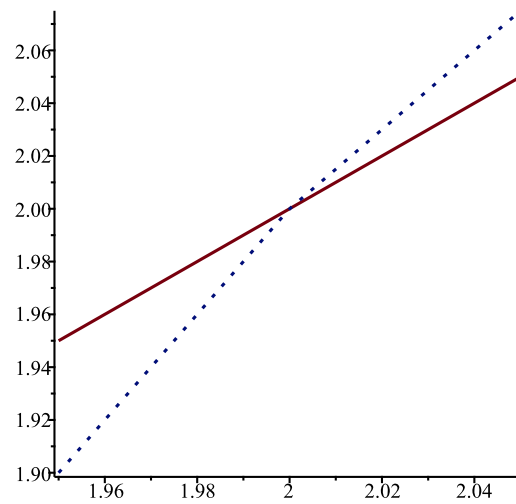
Figure 1. Functions $G(1)$ (solid line) and $H(1)$ when $\lambda \in [1.95, 2.05]$.

Similarly, $G(10)$ and $H(10)$ are both approximately equal to 11.576 when $\lambda \approx 2.1576$; see Figure 2.
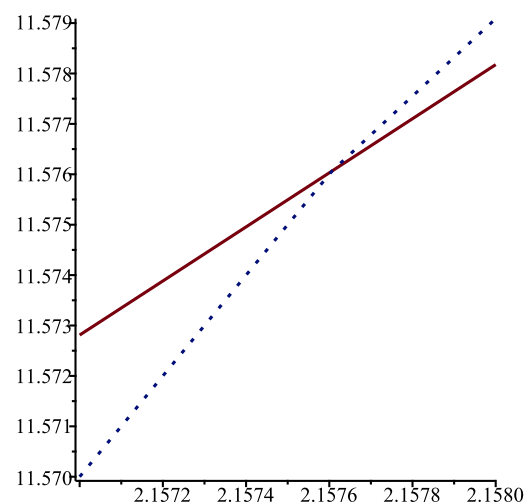
Figure 2. Functions $G(10)$ (solid line) and $H(10)$ when $\lambda \in [2.157, 2.158]$.

4.2. Solution of the Non-Linear Difference Equation

As we mentioned in Section 2, solving the non-linear third-order difference Equation (18), subject to the boundary condition (19), is not an easy task.

We can try to solve Equation (18) recursively. First, making use of the boundary condition (19), we can write that

$$F^2(1) - F(1)(3\lambda - 2) + 2\lambda(\lambda - 1) = 0. \tag{33}$$

It follows that

$$
\begin{aligned}
F(1) &= \frac{1}{2}\left\{(3\lambda - 2) \pm \left[\sqrt{(3\lambda - 2)^2 - 8[\lambda(\lambda - 1)]}\right]\right\} \\
&= \frac{1}{2}\left\{(3\lambda - 2) \pm \sqrt{\lambda^2 - 4\lambda + 4}\right\} \\
&= \frac{1}{2}\left\{(3\lambda - 2) \pm |\lambda - 2|\right\}. \tag{34}
\end{aligned}
$$

Because $F(1)$ can take any real value, it is not obvious to decide which sign to choose in the above equation. If we look at the value of $F(1)$ in Tables 1–4, we see that we must in fact choose the *minus* sign, so that

$$F(1) = \begin{cases} 2(\lambda - 1) & \text{if } \lambda \in (0,2], \\ \lambda & \text{if } \lambda \geq 2. \end{cases} \tag{35}$$

Next, we could use the above expression for $F(1)$ in Equation (18) to determine the value of $F(2)$, and so forth. It is clear that this technique is tedious. Moreover, without having computed the value of $F(x)$ as we did above, determining the right sign to choose in the expression for $F(x)$ that corresponds to the one in Equation (35) is not straightforward.

From what precedes, we may conclude that the method used in this paper to compute the value function explicitly is of interest in itself and could be used to solve boundary value problems for non-linear difference equations, if we can express the function of interest as a particular value function in a stochastic control problem.

To conclude, we will check that the values of the function $F(x)$ given in Table 1 are such that Equation (18), together with the boundary condition (19), is indeed satisfied for $x = 1, 2, \ldots, 10$ when $\lambda = 1/2$.

First, we saw in Equation (35) that $F(1) = -1$ is a solution of the equation if $\lambda = 1/2$. Next, since $F(2)$ is also equal to -1, we must have

$$2(-1)^2 - (-1)[6(1/2) - 4 + (-1) + 0 + 0] + 0 + 2[(1/2) - 1][(-1) + 0]$$
$$+4(1/2)[(1/2) - 1] = 0, \tag{36}$$

which is indeed correct.

The value of $F(3)$ in Table 1 is -2. We must check that

$$2(-2)^2 - (-2)[6(1/2) - 4 + (-1) + 2(-1) + 0] + (-1)[(-1) + 0 + 2(1/2)]$$
$$+2[(1/2) - 1][(-1) + 0] + 4(1/2)[(1/2) - 1] = 0. \tag{37}$$

Again, the result is correct. We can proceed in the same way to check the remaining results. For $x = 10$, we obtain that

$$2(-5)^2 - (-5)[6(1/2) - 4 + (-5) + 2(-4) + (-4)] + (-4)[(-5) + (-4) + 2(1/2)]$$
$$+2[(1/2) - 1][(-5) + (-4)] + 4(1/2)[(1/2) - 1] = 0. \tag{38}$$

5. Conclusions

Homing problems are generally considered for diffusion processes. The author and Kounta [11] extended these problems to the discrete-time case. In [10], the author improved the results found in [11] by finding an explicit expression for the value function.

In the current paper, a homing problem with two optimizers has been defined and solved explicitly. The problem can be interpreted as a stochastic difference game, as one optimizer is trying to minimize the time spent by the controlled stochastic process in the continuation region C, while the second one seeks to maximize the survival time in C.

In Section 2, the equation satisfied by the value function has been derived. This equation is a non-linear third-order difference equation, which is obviously very difficult to solve explicitly.

The technique that we have used in the paper enables us to obtain an exact expression for the solution of a boundary value problem for a non-linear difference equation. This result is of interest in itself.

For the sake of simplicity, we have considered a symmetric random walk, and it has been assumed that the control variables can take only two values. We could of course generalize the results that have been presented in the paper. However, if there are many possible values for the control variables, obtaining an explicit solution to the homing problem considered can be quite tedious.

Funding: This research was supported by the Natural Sciences and Engineering Research Council of Canada. The author also wishes to thank the anonymous reviewers of this paper for their constructive comments.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Hämäläinen, R.P. Nash and Stackelberg solutions to general linear-quadratic two player difference games. I: Open-loop and feedback strategies. *Kybernetika* **1978**, *14*, 38–56.
2. Reddy, P.V.; Zaccour, G. Feedback Nash equilibria in linear-quadratic difference games with constraints. *IEEE Trans. Automat. Control* **2017**, *62*, 590–604. [CrossRef]
3. Lin, Y. Feedback Stackelberg strategies for the discrete-time mean-field stochastic systems in infinite horizon. *J. Frank. Inst.* **2019**, *356*, 5222–5239. [CrossRef]
4. Liu, X.; Liu, R.; Li, Y. Infinite time linear quadratic Stackelberg game problem for unknown stochastic discrete-time systems via adaptive dynamic programming approach. *Asian J. Control* **2020**, *2*, 937–948. [CrossRef]
5. Ju, P.; Li, X.; Lei, J.; Li, T. Optimal linear strategy for stochastic games with one-side control-sharing information. *Int. J. Control* **2022**, 1–7. [CrossRef]
6. Whittle, P. *Optimization over Time*; Wiley: Chichester, UK, 1982; Volume I.
7. Whittle, P. *Risk-Sensitive Optimal Control*; Wiley: Chichester, UK, 1990.
8. Kuhn, J. The risk-sensitive homing problem. *J. Appl. Probab.* **1985**, *22*, 796–803. [CrossRef]
9. Makasu, C. Risk-sensitive control for a class of homing problems. *Automatica* **2009**, *45*, 2454–2455. [CrossRef]
10. Lefebvre, M. An explicit solution to a discrete-time stochastic optimal control problem. *WSEAS Trans. Syst.* **2023**, *22*, Art. #40, 368–371. [CrossRef]
11. Lefebvre, M.; Kounta, M. Discrete homing problems. *Arch. Control Sci.* **2013**, *23*. [CrossRef]
12. Lefebvre, M. The homing problem for autoregressive processes. *IMA J. Math. Control Inform.* **2022**, *39*, 322–344. [CrossRef]
13. Kounta, M.; Dawson, N.J. Linear quadratic Gaussian homing for Markov processes with regime switching and applications to controlled population growth/decay. *Methodol. Comput. Appl. Probab.* **2021**, *23*, 1155–1172. [CrossRef]
14. Feller, W. *An Introduction to Probability Theory and its Applications*; Wiley: New York, NY, USA, 1968; Volume I.