

An Optimal Solution to the Linear Search Problem for a Robot with Dynamics

Irene Ruano De Pablo

Aaron Becker

Timothy Bretl

Abstract—In this paper we derive the control policy that minimizes the total expected time for a point mass with bounded acceleration, starting from the origin at rest, to find and return to an unknown target that is distributed uniformly on the unit interval. We apply our result to proof-of-concept hardware experiments with a planar robot arm searching for a metal object using an inductive proximity sensor. In particular, we show that our approach easily extends to optimal search along arbitrary curves, such as raster-scan patterns that might be useful in other applications like robot search-and-rescue.

I. INTRODUCTION

The classical *linear search problem*, originally posed by Bellman [1] and Beck [2], is the following:

A target z is placed somewhere on the real line \mathbb{R} according to a known probability distribution $g(z)$.

We can search for this target by starting at the origin and moving in either direction at unit speed.

If we recognize the target when we pass it, what policy minimizes the expected time to do so?

In some cases the solution to this problem is intuitively clear. For example, assume that $g(z)$ describes a uniform distribution on the interval $[-a, b]$ for positive constants $a, b > 0$. If $a > b$, we should move first to b and then to a . If $b > a$, we should do the opposite. In other cases, for more general distributions $g(z)$, the situation is much more complex. As a result, this problem has prompted a long string of papers over the past fifty years, some of which have considered extensions like achieving rendezvous with two searchers [3] or doing linear search on a network [4]. Linear search falls within a broader class of search problems that are reviewed by [5], [6].

In this paper we consider a variant of the linear search problem in which the searcher is a point mass that—instead of moving at unit speed—is subject to bounded acceleration. Our goal in this case is to not only find the target but also return to its location in minimum expected time. Although we will assume that the target is distributed uniformly over the unit interval, the optimal policy is no longer intuitively clear. In particular, there is a tradeoff now between *finding* the target (for which faster is better) and *returning* to the target (for which slower is better).

This variant of the linear search problem is of particular relevance to robotics, where the point mass could represent an unmanned aircraft doing search and rescue, the tip of a confocal microscope looking for tagged fluorescent particles, or the end effector of a robot arm searching for objects using a proximity sensor. This variant has also received little previous attention. Perhaps most closely related, the work of Demaine [7] and earlier of Lopez-Ortiz [8] considered the linear search problem with an additional cost for turning, but still assumed unit speed between turns.

Much more general search problems than ours have been considered over the past decade by the robotics community. This work tends to focus on pursuit-evasion games [9], where both the pursuer (the searcher) and the evader (the target) are moving. It addresses complications like probabilistic motion models, probabilistic sensor models, multiple pursuers or evaders, and visibility constraints (e.g., [10]–[16]). It is closely related both to the problem of coverage [17] and to the problem of exploration, either for active localization or mapping [18]. This work may be classified more broadly as planning under uncertainty, for which general solution approaches exist (e.g., POMDP solvers [19]–[21]). However, these approaches rarely lead to exact solutions, and are often computationally prohibitive. Common heuristics include finite horizon policies and policies based on the certainty equivalence principle (i.e., on an assumed decoupling between optimal estimation and optimal control).

In contrast, our linear search problem—although a special case—admits an exact analytical solution, which we will proceed to derive as follows. In Section II, we will formulate our problem as a time-invariant optimal control problem with free terminal time and additive cost. In Section III, we will establish necessary conditions for optimality using the minimum principle [22]–[24], and establish sufficient conditions by explicit computation. The resulting optimal control policy accelerates at the maximum rate for a distance $3 - 2\sqrt{2} \approx 0.17$, then decelerates at a lower but still constant rate until returning to rest at the end of the unit interval. This result corresponds neither to the trajectory that crosses the interval and returns to rest in minimum time, nor to the trajectory that would have been derived from the certainty equivalence principle. In Section IV, we will apply our result to hardware experiments with a planar robot arm searching for a metal object using an inductive proximity sensor. In particular, we will show that our approach easily extends to search along arbitrary curves, such as raster-scan patterns, that might be useful in practical applications. Finally, in Section V, we will consider possibilities for future work.

I. Ruano De Pablo is with the Departamento de Automática y Electrónica Industrial, Universidad Pontificia Comillas, Madrid 28400, España ire.ruanodp@gmail.com

A. Becker is with the Department of Electrical and Computer Engineering and T. Bretl is with the Department of Aerospace Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA {abecker5, tbretl}@illinois.edu

II. PROBLEM STATEMENT

Consider a unit mass with position $q \in \mathbb{R}$ being driven by a force $u \in [-1, 1]$ with dynamics $\ddot{q} = u$. The target is some point $z \in [0, 1]$. All possible target positions are equally likely. The mass starts at $q(0) = 0$. We have a sensor that tells us when we have passed the target, in other words when $q(t) = z$. Our goal is to stop at the target in minimum expected time, in other words to reach $(q, \dot{q}) = (z, 0)$.

A. The Cost to Reach a Particular Target

Define the state as $x = (x_1, x_2)$, where $x_1 = q$ and $x_2 = \dot{q}$. Consider the instant t_0 at which the mass passes the target, so $x_1(t_0) = z$. For convenience, we denote the velocity of the mass at this instant by $v = x_2(t_0)$. After this instant, the time-optimal policy is a bang-bang solution of the form

$$u(t) = \begin{cases} -1 & t \in [t_0, t_0 + t_1] \\ 1 & t \in (t_0 + t_1, t_0 + t_2] \end{cases}$$

for some $t_2 \geq t_1 \geq 0$. To reach $x_1(t_0 + t_2) = x_1(t_0) = z$ and $x_2(t_0 + t_2) = 0$, the following two equations must hold:

$$z = z + \left(vt_1 - \frac{t_1^2}{2} \right) + \left((v - t_1)(t_2 - t_1) + \frac{(t_2 - t_1)^2}{2} \right)$$

$$0 = (v - t_1) + (t_2 - t_1).$$

The first is an expression for the position at time $t_0 + t_2$, and the second is an expression for the velocity at this time. Noting that $z \geq 0$ implies $v \geq 0$, we find that

$$t_1 = \left(1 + \frac{1}{\sqrt{2}} \right) v \quad \text{and} \quad t_2 = \left(1 + \sqrt{2} \right) v.$$

So, the cost of reaching the target after passing it is a linear function of the velocity at which it is passed:

$$t_{\text{return}}(v) = av,$$

where $a = 1 + \sqrt{2}$. The total cost is therefore

$$t_0 + ax_2(t_0).$$

B. The Expected Cost to Reach an Unknown Target

For now, we will assume that $x_2 \geq 0$ until the target is passed. It will turn out that this assumption holds for any optimal policy (Section III-E). Given that z is uniformly distributed over $[0, 1]$, the expected time to reach the target may therefore be found by integration

$$J = \int_0^1 (t + ax_2) dx_1,$$

where t is the time at which the target is passed, i.e., at which $x_1(t) = z$. Since $dx_1 = x_2 dt$, then equivalently

$$J = \int_0^T (t + ax_2) x_2 dt,$$

where T is the free final time at which $x_1(T) = 1$. For convenience, we will define an additional state $x_3 = t$, so that $\dot{x}_3 = 1$, in order to make the system time-invariant. The total expected cost becomes

$$J = \int_0^T (x_2 x_3 + ax_2^2) dt.$$

C. The Resulting Optimal Control Problem

Our goal is to minimize

$$J = \int_0^T (x_2 x_3 + ax_2^2) dt \quad (1)$$

for free final time T subject to the dynamics

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u, \quad \text{and} \quad \dot{x}_3 = 1,$$

the constraints

$$u \in [-1, 1] \quad \text{and} \quad x_2 \geq 0,$$

and the boundary conditions

$$x_1(0) = 0, \quad x_2(0) = 0, \quad x_3(0) = 0, \quad \text{and} \quad x_1(T) = 1,$$

where $a = 1 + \sqrt{2}$. The Hamiltonian is

$$H(x, p, u) = p_0 (ax_2^2 + x_2 x_3) + p_1 x_2 + p_2 u + p_3,$$

where $p_0 \geq 0$ is a constant. The adjoint equations are

$$\begin{aligned} \dot{p}_1 &= -\frac{\partial H}{\partial x_1} = 0 \\ \dot{p}_2 &= -\frac{\partial H}{\partial x_2} = -p_0 (2ax_2 + x_3) - p_1 \\ \dot{p}_3 &= -\frac{\partial H}{\partial x_3} = -p_0 x_2. \end{aligned} \quad (2)$$

The minimum principle tells us that any optimal policy u^* must satisfy

$$0 = H^*(x^*, p_0^*, p^*, u^*) \leq \min_u H(x, p_0, p, u) \quad (3)$$

along the optimal trajectory x^* for some p_0^* and p^* , not both zero [22]–[24].

III. SOLUTION APPROACH

In Sections III-A through III-C, we will establish necessary conditions for optimality. Our main result (Lemmas 3.2 and 3.3) will be to identify five candidate control policies, each of which can be parameterized by no more than two switching times. Then, in Section III-D, we will eliminate all but one of these candidate policies by minimizing (1) with respect to the switching times. Finally, in Section III-E, we will relax the assumption $x_2 \geq 0$ made in our problem statement, in particular showing that an optimal trajectory must satisfy this constraint. We will conclude in Section III-F with a comparison between the optimal policy that we derive and two reasonable heuristic policies.

A. Valid Inputs

Lemma 3.1: An optimal policy must satisfy

$$u(t) \in \left\{ -1, -\frac{1}{2a}, 1 \right\}$$

for (almost) all $t \in [0, T]$.

Proof: Consider a particular time $t_1 \in [0, T]$. If $p_2(t_1) \neq 0$, then the condition (3) tells us that

$$u(t_1) = -\text{sign } p_2(t_1),$$

in other words that $u(t_1) \in \{-1, 1\}$. If $p_2(t_1) = 0$, then the condition (3) provides no information. First, assume there exists no interval $[t_1, t_2) \subset [0, T)$ such that $p_2(t) = 0$ for all $t \in [t_1, t_2)$. In this case, the time t_1 at which $p_2(t_1) = 0$ is isolated, so since $u(t_1)$ is bounded it may be safely ignored. Conversely, assume the existence of such an interval $[t_1, t_2)$, within which we must have $\dot{p}_2(t) = 0$ and hence also $\ddot{p}_2(t_1) = 0$. If $p_0 = 0$, then this condition implies that $p_1 = 0$. But, since $H = 0$ along an optimal trajectory, we conclude that $p_3(t_1) = 0$ as well, hence that $p_3 = 0$ always. Since not both of p_0 and p can vanish, we must have $p_0 > 0$, hence we may assume without loss of generality that $p_0 = 1$. So, in this case, we must have

$$u(t_1) = -\frac{1}{2a},$$

and therefore we have our result. \blacksquare

In fact, it will turn out that an optimal policy is a finite sequence of these inputs. For convenience, we will label intervals along which $u = 1$ as \uparrow , intervals along which $u = -1$ as \downarrow , and intervals along which $u = -1/2a$ as \curvearrowright .

B. Normal Extremals

Lemma 3.2: An optimal policy for which $p_0 > 0$ must be of type \uparrow , $\uparrow\downarrow$, $\uparrow\curvearrowright$, $\uparrow\curvearrowright\uparrow$, or $\uparrow\curvearrowright\downarrow$.

Proof: We assume without loss of generality that $p_0 = 0$. To satisfy $x_2 \geq 0$, the condition (3) implies that we must have $p_2(0) < 0$, hence that an optimal policy must begin with an interval of type \uparrow . For convenience, we will denote the initial condition by $p_2(0) = -c$ and the interval by $[0, t_1]$, where $c > 0$ and $t_1 > 0$. For all $t \in [0, t_1]$, we have

$$\dot{x}_2(t) = 1 \quad \Rightarrow \quad x_2(t) = t$$

and so

$$\dot{p}_2(t) = -(1 + 2a)t - p_1.$$

Integrating, we find

$$p_2(t) = -c - \left(\frac{1 + 2a}{2}\right)t^2 - p_1 t.$$

If

$$p_1 > -\sqrt{2(1 + 2a)c},$$

then there is no time $t_1 > 0$ at which $p_2(t_1) = 0$, hence the entire policy is of type \uparrow . Alternatively, if

$$p_1 < -\sqrt{2(1 + 2a)c},$$

then there exists $t_1 > 0$ at which $p_2(t_1) = 0$, and at the first such time we have

$$\dot{p}_2(t_1) = \sqrt{p_1^2 - 2(1 + 2a)c} > 0$$

regardless of $u(t_1)$, hence we switch to an interval of type \downarrow for which

$$\dot{x}_2(t) = -1 \quad \Rightarrow \quad x_2(t) = x_2(t_1) - (t - t_1)$$

and so

$$\dot{p}_2(t) = \dot{p}_2(t_1) + (2a - 1)(t - t_1)$$

when $t \geq t_1$. Since $\dot{p}_2(t_1) > 0$ by assumption and $2a - 1 = 1 + 2\sqrt{2} > 0$, then $\dot{p}_2(t) > 0$ for all $t > t_1$. So, the entire policy in this case is of type $\uparrow\downarrow$. Finally, if

$$p_1 = -\sqrt{2(1 + 2a)c},$$

then there exists $t_1 > 0$ at which $p_2(t_1) = \dot{p}_2(t_1) = 0$. Three subsequent policies satisfy the minimum principle:

$$\begin{aligned} u(t) &= \begin{cases} -1/2a & t_1 \leq t \end{cases} \\ &\Rightarrow \dot{p}_2(t) = \begin{cases} 0 \end{cases} & \Rightarrow \curvearrowright \\ u(t) &= \begin{cases} -1/2a & t_1 \leq t \leq t_2 \\ 1 & t_2 < t \end{cases} \\ &\Rightarrow \dot{p}_2(t) = \begin{cases} 0 \\ -(2a + 1)(t - t_2) \end{cases} & \Rightarrow \curvearrowright\uparrow \\ u(t) &= \begin{cases} -1/2a & t_1 \leq t \leq t_2 \\ -1 & t_2 < t \end{cases} \\ &\Rightarrow \dot{p}_2(t) = \begin{cases} 0 \\ (2a - 1)(t - t_2) \end{cases} & \Rightarrow \curvearrowright\downarrow \end{aligned}$$

The entire policy in each case becomes $\uparrow\curvearrowright$, $\uparrow\curvearrowright\uparrow$, and $\uparrow\curvearrowright\downarrow$, respectively. And so, we have our result. \blacksquare

C. Abnormal Extremals

Lemma 3.3: An optimal policy for which $p_0 = 0$ must be of type \uparrow or $\uparrow\downarrow$.

Proof: If $p_0 = 0$ then the Hamiltonian becomes

$$H(x, p, u) = p_1 x_2 + p_2 u + p_3 = 0,$$

where both p_1 and p_3 are now constant. In particular, the minimum principle requires that

$$p_1 x_2(0) + p_2(0)u(0) + p_3 = 0. \quad (4)$$

Consider the case for which $p_2(0) = \dot{p}_2(0) = 0$. Equation (2) requires that $p_1 = 0$, and Eq. (4) then requires that $p_3 = 0$. But, we cannot have both p_0 and p vanish. It must therefore be the case that either $p_2(0) \neq 0$ or $\dot{p}_2(0) \neq 0$, or both.

First, assume that $\dot{p}_2(0) = 0$ so that $p_2(0) \neq 0$. From (2), we have $p_1 = 0$ and so $p_2(t) = p_2(0)$ for all t . Since we must have $x_2 \geq 0$, the condition (3) implies that $p_2(0) > 0$ and so the entire policy is of type \uparrow .

Now, assume that $\dot{p}_2 \neq 0$ so that $p_1 \neq 0$. Then, we integrate to find $p_2(t) = p_2(0) - p_1 t$. This expression is linear in t , so there is at most one switch. Given the constraint $x_2 \geq 0$, the resulting policy must be either of type \uparrow or of type $\uparrow\downarrow$. \blacksquare

D. Finding the Optimal Policy

Lemma 3.4: The optimal policy is of type $\uparrow\curvearrowright$ and can be expressed as

$$u^*(t) = \begin{cases} 1 & 0 \leq t < 2 - \sqrt{2} \\ -1/2a & 2 - \sqrt{2} \leq t < 2 + \sqrt{2} \end{cases}$$

or equivalently as

$$u^*(x_1) = \begin{cases} 1 & 0 \leq x_1 < 2 - \sqrt{2} \\ -1/2a & 2 - \sqrt{2} \leq x_1 < 1 \end{cases}.$$

The corresponding cost is

$$J^* = \frac{2}{3} (2 + \sqrt{2}) = \frac{2T}{3},$$

where T is the final time, i.e., the time at which $x_1(T) = 1$.

Proof: Lemmas 3.2-3.3 imply that the optimal policy must be of type \uparrow , $\uparrow\downarrow$, $\uparrow\curvearrowright$, $\uparrow\curvearrowleft\uparrow$, or $\uparrow\curvearrowleft\downarrow$, so in any case

$$u(t) = \begin{cases} 1 & 0 \leq t < t_1 \\ -1/2a & t_1 \leq t < t_1 + t_2 \\ -1 & t_1 + t_2 \leq t < t_1 + t_2 + t_3 \end{cases}, \quad (5)$$

where $t_1, t_2, t_3 \geq 0$ and $t_1 + t_2 + t_3 = T$. We want to find the values of t_1 , t_2 , and t_3 that minimize (1), and to establish the corresponding cost. We do this as follows:

- Integrate (5) to find $x_1(t)$ and $x_2(t)$ as functions of t_1 , t_2 , and t_3 , given $x_1(0) = x_2(0) = 0$.
- Apply the final conditions

$$x_1(t_1 + t_2 + t_3) = 1, \quad x_2(t_1 + t_2 + t_3) = v$$

for arbitrary $v \geq 0$ to eliminate t_2 and t_3 , leaving our expressions for $x_1(t)$ and $x_2(t)$ in terms of the parameters t_1 and v only.

- Establish bounds on t_1 as a function of v . In particular, we note that t_1 is minimized when $t_3 = 0$ and maximized when $t_2 = 0$, resulting in the bounds

$$\sqrt{(6 - 4\sqrt{2}) + 2(-1 + \sqrt{2})v^2} \leq t_1 \leq \sqrt{1 + \frac{v^2}{2}}, \quad (6)$$

where $0 \leq v \leq \sqrt{2}$.

- Plug in $x_2(t)$ and $T = t_1 + t_2 + t_3$ to find the cost J , given by (1), as a function of t_1 and v . By evaluating $\partial J / \partial t_1$ and ignoring solutions to $\partial J / \partial t_1 = 0$ for which $t_1 < 0$, we establish that candidate extremals of J occur at

$$t_1 = 2 - \sqrt{2} \quad \text{and} \quad t_1 = \sqrt{1 + \frac{v^2}{2}}.$$

Comparing these candidates with the bounds (6), we note that

$$\begin{aligned} & \left[\sqrt{(6 - 4\sqrt{2}) + 2(-1 + \sqrt{2})v^2}, \sqrt{1 + \frac{v^2}{2}} \right] \\ & \subseteq \left[2 - \sqrt{2}, \sqrt{1 + \frac{v^2}{2}} \right] \end{aligned}$$

for all $0 \leq v \leq \sqrt{2}$. Furthermore, it is easy to verify that

$$\frac{\partial^2 J}{\partial t_1^2} \Big|_{t_1 = \sqrt{1 + \frac{v^2}{2}}} < 0$$

for all $0 \leq v \leq \sqrt{2}$. As a consequence, the minimum value of J occurs at

$$t_1^* = \sqrt{(6 - 4\sqrt{2}) + 2(-1 + \sqrt{2})v^2}.$$

- Plug in t_1^* to find J as a function of v only. By evaluating $\partial J / \partial v$, we establish that candidate extremals occur at $v = 0$ and $v = \sqrt{2}$. We find that

$$\frac{\partial^2 J}{\partial v^2} \Big|_{v=0} = 0 \quad \text{and} \quad \frac{\partial^2 J}{\partial v^2} \Big|_{v=\sqrt{2}} = -2\sqrt{2} < 0.$$

We immediately conclude that J is minimum at

$$v^* = 0, \quad t_1^* = \sqrt{6 - 4\sqrt{2}} = 2 - \sqrt{2}.$$

Note that, for these values, we recover

$$t_2^* = 2\sqrt{2}, \quad t_3^* = 0.$$

As a consequence, the optimal policy is of type $\uparrow\curvearrowright$ and can be expressed as

$$u^*(t) = \begin{cases} 1 & 0 \leq t < 2 - \sqrt{2} \\ -1/2a & 2 - \sqrt{2} \leq t < 2 + \sqrt{2} \end{cases}$$

or equivalently as

$$u^*(x_1) = \begin{cases} 1 & 0 \leq x_1 < 3 - 2\sqrt{2} \\ -1/2a & 3 - 2\sqrt{2} \leq x_1 < 1 \end{cases}.$$

The corresponding cost is $J^* = \frac{2}{3} (2 + \sqrt{2})$.

In exactly the same way, we can verify that the only remaining candidate policy $\uparrow\curvearrowleft\uparrow$, expressed as

$$u(t) = \begin{cases} 1 & 0 \leq t < t_1 \\ -1/2a & t_1 \leq t < t_1 + t_2 \\ -1 & t_1 + t_2 \leq t < t_1 + t_2 + t_3 \end{cases},$$

is optimal for the same choice of parameters

$$t_1^* = 2 - \sqrt{2}, \quad t_2^* = 2\sqrt{2}, \quad t_3^* = 0,$$

hence that it also reduces to the same policy $\uparrow\curvearrowright$. ■

E. Relaxing the Constraint $x_2 \geq 0$

We have assumed that $x_2 \geq 0$, i.e., that the velocity must be non-negative always. If we relax this assumption, it is still possible to show that $x_2 \geq 0$ along any optimal trajectory, hence that the optimal policy remains as we computed it in Section III-D. We will sketch a proof in this section, omitting the details. We will rely on the general principle that any subset of an optimal trajectory must, itself, also be optimal.

Assume that $x_2(t) < 0$ for some $t > 0$. Then, there must exist some time $t_1 \geq 0$ at which $x_2(t_1) = 0$, and furthermore some time $t_2 > t_1$ satisfying $x_1(t_2) = x_1(t_1)$. Denote the velocity at time t_2 by $x_2(t_2) = v$, where we may assume without loss of generality that $v \geq 0$. It is easy to verify that the optimal policy on the interval $[t_1, t_2]$, i.e., the policy that minimizes the time required to transition from $(x_1, x_2) = (0, 0)$ to $(x_1, x_2) = (0, v)$, is

$$u(t) = \begin{cases} -1 & t_1 \leq t < t_1 + v(\sqrt{2}/2) \\ 1 & t_1 + v(\sqrt{2}/2) \leq t < t_1 + v(1 + (\sqrt{2}/2)) \end{cases}.$$

The resulting cost is $t_2 - t_1 = (1 + \sqrt{2})v = av$. In fact, we see that the effect of allowing $x_2(t) < 0$ is to allow points of

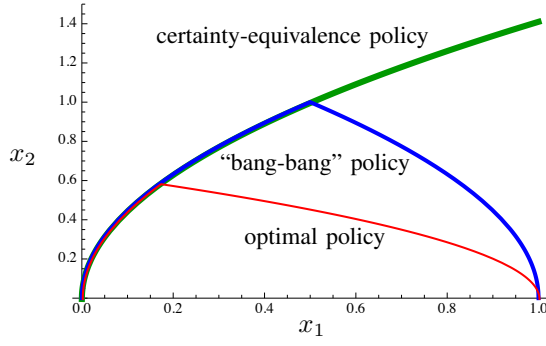


Fig. 1. Velocity profile for the optimal policy (red), the “bang-bang” policy (blue), and the policy that would result from the application of a certainty-equivalence principle (green), i.e., of decoupling estimation and control.

discontinuity t at which the velocity jumps from $x_2(t) = 0$ to $x_2(t) = v \geq 0$, at a cost of av . In other words, we can describe any trajectory on the interval $[0, T]$ as a sequence of shorter trajectories on the subintervals

$$[t_0, t_1], [t_1, t_2], \dots, [t_{n-1}, t_n].$$

For each subinterval $i = 1, \dots, n$, we may choose the initial velocity $x_2(t_{i-1}) = v_{i-1}$ and require that the final velocity is zero. Also, within each subinterval, we may assume $x_2 \geq 0$. We note that any optimal trajectory must also be optimal when restricted to any of its subintervals. This decomposition suggests the following strategy of proof:

- Repeat the above analysis but for arbitrary initial velocity $x_2(0) = v_0$ and for modified cost $J' = av_0 + J$.
- Show that the optimal policy satisfies $v_0^* = 0$, i.e., it is always best to begin each subinterval at zero velocity.
- Conclude that our assumption $x_2 \geq 0$ was valid, hence that we recover the same optimal policy.

The technical details are not hard, just tedious. For example, we must consider several additional candidate policies (e.g., $\downarrow\curvearrowright\uparrow$ and $\downarrow\curvearrowleft\downarrow$) and we must optimize over three variables (t_1, v, v_0) instead of two (t_1, v) .

F. Comparison with Heuristic Strategies

Figure 1 shows the optimal velocity profile as compared to two other alternative strategies that may at first seem reasonable. The first alternative—a “bang-bang” strategy—crosses the interval and returns to rest in minimum time. It is easy to verify that this strategy incurs a cost that is about 15% higher than optimal. The second alternative—a certainty equivalence strategy—continues to accelerate all the way across the interval. It is again easy to verify that this strategy incurs a cost that is about 41% higher than optimal.

Why do we call this second alternative a “certainty equivalence” strategy? Note that, having moved a distance $x_1 \in [0, 1)$, we know the target is uniformly distributed on the interval $[x_1, 1)$. The mean of this distribution—a common choice of best estimate—is at the position

$$\frac{x_1 + 1}{2} > x_1.$$

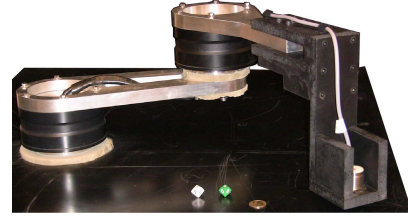


Fig. 2. Two-link arm (lengths 0.3m and 0.45m) used in our experiments. The end-effector is an inductive proximity sensor that detects metal objects.

The time-optimal control policy to reach this position—assuming that it is, indeed, the location of the target—is a “bang-bang” policy that accelerates at maximum rate until reaching the halfway point $(x_1 + 3)/4$. However, this halfway point will recede as we move. So, if we assume for the purposes of computing the optimal control that our best estimate of the target position is correct—i.e., if we apply what is called the *certainty equivalence principle*, a common heuristic when dealing with more general search problems—then we do exactly the wrong thing and never stop accelerating.

IV. HARDWARE EXPERIMENTS

To validate our solution approach, we applied our results to hardware experiments with a two-link planar robot arm (Fig. 2). Each link was powered by a direct-drive brushless DC servo motor with encoder feedback. Planning and control were done on an external PC with a 1kHz control loop. The end-effector carried an inductive proximity sensor that detected metal objects within a radius of 15mm but that did not respond to nonmetallic objects. In our experiments we used a US \$1 coin, placed at unknown locations.

First, we considered a straight-line search path of length 1m (see the video attachment). We used task-space inverse dynamics to generate reference torques for each joint and a computed torque method to control each motor [25]. As a consequence, by defining conservative acceleration bounds in the task-space (i.e., on the motion of the end-effector), we could model the robot exactly as described in Section II. Figure 3 shows example velocity profiles for both the optimal policy and for the alternative “bang-bang” policy along with aggregate results for the optimal policy and for both of the alternatives we considered in Section III-F. These results match the theory developed in Sections II-III.

A natural extension of our work would use space filling curves to search two dimensional areas. As a proof-of-concept, we considered the raster scan pattern in Fig. 4. Although the search path is now a smooth curve, the result is still a linear search problem, and so can be addressed with our solution approach. The only difference is the introduction of configuration-dependent velocity constraints, in particular at the switch-back. Although we do not prove it here, these constraints are easily handled within the same framework.

V. CONCLUSION

We presented an optimal control policy that minimizes the total expected time for a point mass with bounded

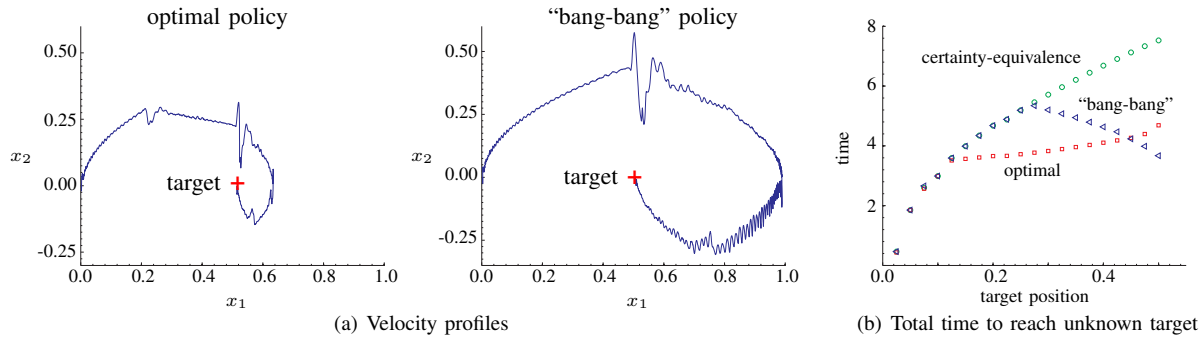


Fig. 3. Experimental results: (a) Optimal and “bang-bang” velocity profiles for one target location. The optimal policy takes longer to detect the target, but returns more quickly. (b) Total time as a function of target position. Each data point is averaged over five trials.

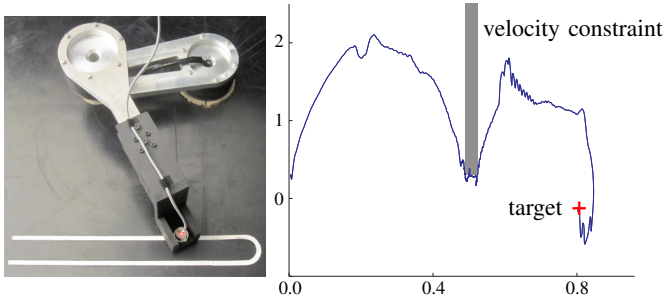


Fig. 4. A raster-scan search path (left) and the corresponding optimal velocity profile (right).

acceleration, starting from the origin at rest, to find and return to an unknown target that is distributed uniformly on the unit interval. We derived this policy using the minimum principle. We applied the result to experiments with a planar robot arm, in particular showing that our “linear search problem” is not confined to straight lines, but rather is easily extended to optimal search along arbitrary curves like raster-scan patterns. Opportunities for future work include extending our results to handle configuration-dependent constraints on velocity and acceleration, to handle target distributions that are non-uniform, and to handle sensor uncertainty. Our results may also simplify the problem of planning optimal raster patterns to handle targets distributed across a surface or volume rather than along a given smooth curve.

VI. ACKNOWLEDGEMENTS

Support was provided by NSF grants CNS-0931871 and CMMI-0956362. Thanks to Dušan Stipanović and to Seth Hutchinson for helpful discussion.

REFERENCES

- [1] R. Bellman, “Problem 63-9, an optimal search,” *SIAM Review*, vol. 5, no. 3, Jul., 1963.
- [2] A. Beck, “On the linear search problem,” *Israel Journal of Mathematics*, vol. 2, no. 4, pp. 221–228, 12 1964.
- [3] S. Alpern and A. Beck, “Pure strategy asymmetric rendezvous on the line with an unknown initial distance,” *Operations Research*, vol. 48, no. 3, pp. 498–501, May - Jun., 2000.
- [4] S. Alpern, V. Baston, and S. Gal, “Searching symmetric networks with utilitarian-postman paths,” *Networks*, vol. 53, no. 4, pp. 392–402, 2009.
- [5] M. G. Monticino, M. G. Monticino, and J. R. Weisinger, “A survey of the search theory literature,” *Naval Research Logistics*, vol. 38, no. 4, pp. 469–494, 1991.
- [6] S. Alpern and S. Gal, *The theory of search games and rendezvous*. Springer, 2003.
- [7] E. D. Demaine, S. P. Fekete, and S. Gal, “Online searching with turn cost,” *Theoretical Computer Science*, vol. 361, no. 2-3, pp. 342–355, 9 2006.
- [8] A. Lopez-Ortiz, “On-line target searching in bounded and unbounded domains,” Ph.D. dissertation, U. Waterloo, Ont., Canada, 1996.
- [9] L. Guibas, J.-C. Latombe, S. M. LaValle, D. Lin, and R. Motwani, “A visibility-based pursuit-evasion problem,” *Int. J. of Computational Geometry and Applications*, vol. 9, no. 4-5, pp. 471–493, 1999.
- [10] R. Vidal, O. Shakernia, H. J. Kim, D. H. Shim, and S. Sastry, “Probabilistic pursuit-evasion games: theory, implementation, and experimental evaluation,” *IEEE Trans. Robot. Autom.*, vol. 18, no. 5, pp. 662–669, Oct. 2002.
- [11] V. Isler, S. Kannan, and S. Khanna, “Randomized pursuit-evasion in a polygonal environment,” *IEEE Trans. Robot.*, vol. 21, no. 5, pp. 875 – 884, oct. 2005.
- [12] F. Bourgault, T. Furukawa, and H. Durrant-Whyte, “Optimal search for a lost target in a bayesian world,” *Field and Service Robotics*, pp. 209–222, 2006.
- [13] S. P. Fekete, R. Klein, and A. Nüchter, “Online searching with an autonomous robot,” *Computational Geometry*, vol. 34, no. 2, pp. 102–115, 5 2006.
- [14] B. P. Gerkey, S. Thrun, and G. Gordon, “Visibility-based pursuit-evasion with limited field of view,” *Int. J. Rob. Res.*, vol. 25, no. 4, pp. 299–315, 4 2006.
- [15] T. H. Chung, “On probabilistic search decisions under searcher motion constraints,” in *WAFR*, 2008.
- [16] C. F. Chung and T. Furukawa, “Coordinated pursuer control using particle filters for autonomous search-and-capture,” *Robotics and Autonomous Systems*, vol. 57, no. 6-7, pp. 700–711, 6 2009.
- [17] H. Choset, “Coverage for robotics—a survey of recent results,” *Annals of Mathematics and Artificial Intelligence*, vol. 31, no. 1, pp. 113–126, 10 2001.
- [18] T. Kollar and N. Roy, “Trajectory optimization using reinforcement learning for map exploration,” *Int. J. Rob. Res.*, vol. 27, no. 2, pp. 175–196, 2 2008.
- [19] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics*. The MIT Press, 2005.
- [20] J. Davidson and S. Hutchinson, “A sampling hyperbelief optimization technique for stochastic systems,” in *WAFR*, 2008.
- [21] H. Kurniawati, D. Hsu, and W. Lee, “SARSOP: Efficient point-based POMDP planning by approximating optimally reachable belief spaces,” in *Robotics: Science and Systems*, 2008.
- [22] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*. John Wiley, 1962.
- [23] M. Athans and P. L. Falb, *Optimal Control: An Introduction to the Theory and Its Applications*. McGraw-Hill, 1966.
- [24] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control*. Hemisphere Publishing, 1975.
- [25] M. W. Spong, S. Hutchinson, and M. Vidyasagar, *Robot Modeling and Control*. John Wiley & Sons, 2006.