or equivalently as

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

The corresponding cost is

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

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 ↑, ↑↓, ↑↷, ↑↷↑, or ↑↷↓, so in any case

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

where $t_1, t_2, t_3 \ge 0$ 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 $v \ge 0$ 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

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

  where $0 \le v \le \sqrt{2}$.
- 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

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

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

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

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

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

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

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

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

  We immediately conclude that J is minimum at

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

  Note that, for these values, we recover

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

  As a consequence, the optimal policy is of type ↑↷ and can be expressed as

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

  or equivalently as

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

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

In exactly the same way, we can verify that the only remaining candidate policy ↑↷↑, expressed as

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

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 ↑↷. ■

E. Relaxing the Constraint $x_2 \ge 0$

We have assumed that $x_2 \ge 0$, i.e., that the velocity must be non-negative always. If we relax this assumption, it is still possible to show that $x_2 \ge 0$ 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 $t_1 \ge 0$ 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 $v \ge 0$. 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

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

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

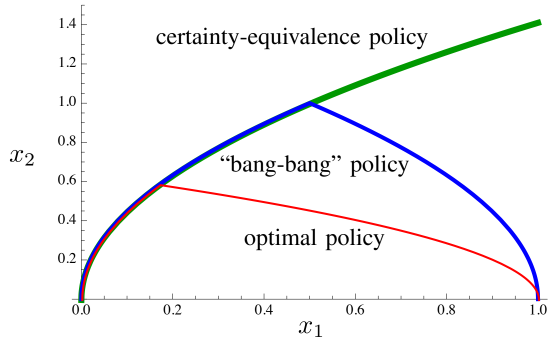


Fig. 1. Velocity profile for the optimal policy (red), the "bang-bang" policy (blue), and the policy that would result from the application of a certainty-equivalence principle (green), i.e., of decoupling estimation and control.

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

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

For each subinterval $i = 1, \ldots, n$, 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 $\downarrow\!\curvearrowright\!\downarrow$) 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

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

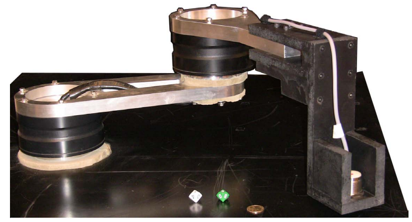


Fig. 2. Two-link arm (lengths 0.3m and 0.45m) used in our experiments. The end-effector is an inductive proximity sensor that detects metal objects.

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

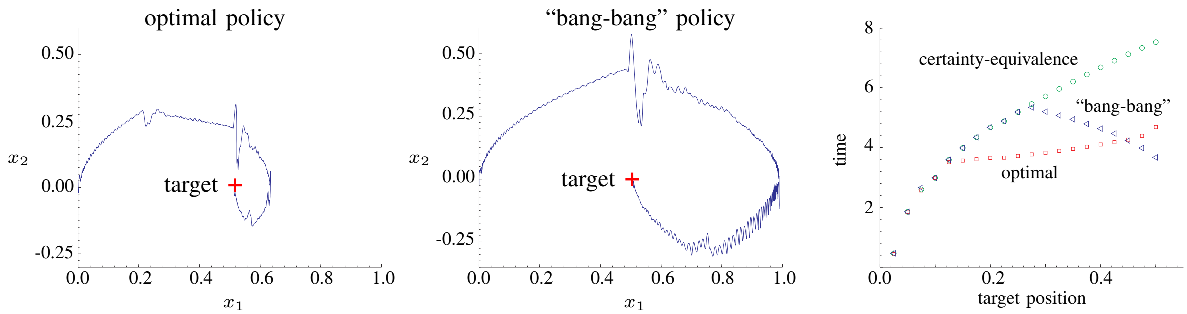
To validate our solution approach, we applied our results to hardware experiments with a two-link planar robot arm (Fig. 2). Each link was powered by a direct-drive brushless DC servo motor with encoder feedback. Planning and control were done on an external PC with a 1kHz control loop. The end-effector 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



(a) Velocity profiles (b) Total time to reach unknown target

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.

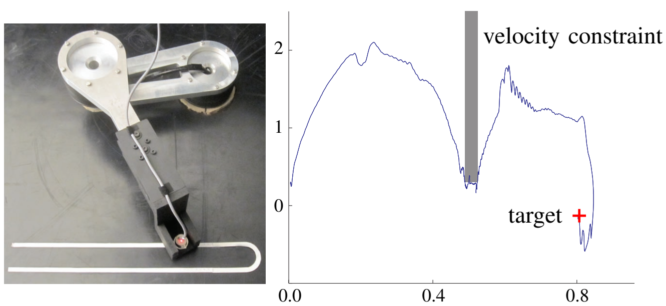


Fig. 4. A raster-scan search path (left) and the corresponding optimal velocity profile (right).

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.