Pseudo Approximation Algorithms with Applications to Optimal Motion Planning

Tetsuo Asano *
School of Information Science, JAIST, Japan
David Kirkpatrick †
University of British Columbia, Canada
Chee Yap ‡
Courant Institute, New York University, USA

June 20, 2005

Abstract

We introduce a technique for computing approximate solutions to optimization problems. If X is the set of feasible solutions, the standard goal of approximation algorithms is to compute $x \in X$ that is an ϵ-**approximate solution** in the following sense:

$$d(x) \leq (1+\varepsilon)d(x^*)$$

where $x^* \in X$ is an optimal solution, $d : X \to \mathbb{R}_{\geq 0}$ is the optimization function to be minimized, and $\varepsilon > 0$ is an input parameter. Our approach is to first devise algorithms that compute **pseudo ε-approximate solutions** satisfying the bound

$$d(x) \leq d(x^*_R) + \varepsilon R$$

where $R > 0$ is a new input parameter. Here x^*_R denotes an optimal solution in the space X_R of R-constrained feasible solutions. The parameter R provides a stratification of X in the sense that (1) $X_R \subseteq X_{R'}$, for $R < R'$ and (2) $X_R = X$ for R sufficiently large.

We first describe a highly efficient scheme for converting a pseudo ε-approximation algorithm into a true ε-approximation algorithm. This scheme is useful because pseudo approximation algorithms seem to be easier to construct than ε-approximation algorithms. Another benefit is that our algorithm is automatically precision-sensitive.

*Support from the Ministry of Education, Science, Sports and Culture, Grant-in-Aid of Japan for Scientific Research.

†Support from the Natural Science and Engineering Research Council of Canada is greatfully acknowledged.

‡Supported by NSF/ITR Grant #CCR-0082056.

We apply our technique to two problems in robotics: (A) Euclidean Shortest Path (3ESP), namely the shortest path for a point robot amidst polyhedral obstacles in 3D, and (B) d_1-optimal motion for a rod moving amidst planar obstacles (1ORM). Previously, no polynomial time ε-approximation algorithm for (B) was known. For (A), our new solution is simpler than previous solutions and has an exponentially smaller complexity in terms of the input precision.

1 Introduction

The design of approximation algorithms is an important theme in the study of optimization problems. The standard goal here is to compute feasible solutions x that are **ε-approximate** in the following sense. Suppose X is the space of feasible solutions, and $d : X \to \mathbb{R}_{\geq 0}$ is the criterion for minimization. Then

$$d(x) \leq (1 + \varepsilon)d(x^*) \tag{1}$$

where x^* is an optimum solution.

Especially in geometric settings, we can often parametrize the space X by a real parameter R to yield subspaces $\{X_R \subseteq X : R \geq 0\}$ with two properties: (1) if $R < R'$ then $X_R \subseteq X_{R'}$, and (2) $X_R = X$ for R sufficiently large. We call R the **search radius**. Let x^*_R denote the optimum solution in X_R. Then a **pseudo ε-approximation algorithm** is one[1] that constructs x satisfying

$$d(x) \leq d(x^*_R) + \varepsilon R \tag{2}$$

for any given $\varepsilon > 0$ and $R \geq 0$. In many situations such algorithms are easier to construct than a true ε-approximation algorithm. (Intuitively, the parameter R serves to offset the complexity attributable to the size of the search space by permitting a larger relative error.)

We will show that, under some fairly natural assumptions on $d(x^*_R)$, we can systematically convert a pseudo ε-approximation algorithm into an efficient ε-approximation algorithm which is automatically **precision-sensitive**. The advantage for this approach derives from (i) the relative ease of constructing pseudo approximation algorithms as compared to approximation algorithms, (ii) the use of the above generic conversion scheme, and (iii) a clearer understanding of those aspects of the approximation process that are sensitive to the precision of the input (as well as the nature of this dependence).

Two NP-Hard Optimum Motion Planning Problems. Approximation algorithms take on a special significance when applied to problems that are provably intractable. We will apply the above technique to derive approximation algorithms for two NP-hard problems in the area of robot motion planning. Although motion planning [1, 2, 9, 11, 24, 26] has been extensively studied since

[1] In combinatorial optimization, the term "pseudo-approximation" sometimes refer to a feasible solution when the original constraints are relaxed.

the early 1980s, very little is known about the problem of shortest length motion. Indeed, the only known efficient general algorithms apply only to the case where robot body is a disc in 2D. This paucity of efficient algorithms is not for lack of interest in optimum motion (see below for an overview of the literature in the case of a rod). In retrospect, we now understand the lack of success:

(A) In 3D, Canny and Reif [6] show that the shortest path for a point robot moving amidst polyhedral obstacles is NP-hard to compute. This problem is known as Euclidean Shortest Path in 3D (3ESP).

(B) In 2D, Asano et al [3] define the "d_1-distance" of a rod motion to be the length of the trajectory of the midpoint of the rod and prove the d_1-optimal motion of a rod amidst polygonal obstacles is NP-hard to compute. Let 1ORM be the acronym for this problem. The rod is a directed line segment, and for now the reader may interpret the "d_1-distance" of a rod motion to be the length of the trajectory of the midpoint of the rod (see Appendix 1 for details).

Both (A) and (B) are the simplest optimum motion planning problems in dimensions 3 and 2, respectively, that go beyond a planar disc robot. These NP-hardness results were unexpected when they were first obtained. In any case, they immediately motivated the search for approximation algorithms. For Problem (A), Papadimitriou [16] gave the first approximation algorithm. This was improved and sharpened by Choi et al [8]. Sellen et al [20] further constructed the first **precision-sensitive** algorithm that constructs the true combinatorially shortest path sequence, as well as an ε-approximate path on this sequence.

For Problem (B), no previous ε-approximation algorithm was known. Asano et al [3] provided an approximation algorithm that, for any $\varepsilon, \varepsilon' > 0$, produces a motion μ for the rod satisfying

$$d_1(\mu) \leq (1+\varepsilon) d_1(\mu^*) + O(n^2 \varepsilon') \tag{3}$$

where μ^* is a d_1-optimal motion. The algorithm runs in time is $O(n^4 \alpha(n) K)$ where $\alpha(n)$ is the inverse Ackermann function and

$$K = K(L, \varepsilon, \varepsilon') = \frac{L - \log \varepsilon'}{\log(1 + \varepsilon L^{-L})}$$

It is assumed that the input description involves only rational numbers, L is the maximum bit length of the input integers, and n is the number of obstacle corners. We will use the same parameters in the algorithms of this paper. All complexity bounds in this paper are in the algebraic complexity model [8].

Unfortunately, this algorithm falls short of being a true ε-approximation algorithm in two respects. First, the analysis assumes a *quantitative* form of their shortest path characterization that, while plausible, has not actually been proved. Second, the bound has an additive term "$+O(n^2\varepsilon')$". which is independent of the main term "$(1+\varepsilon) d_1(\mu^*)$". Since ε' can be set as small as we like (at some increase in the running time), we will obtain a ϵ-approximation algorithm if we could choose it so that the main term dominates the additive

term. To do this, we need some á priori lower bound on $d_1(\mu^*)$ and then choose ε' accordingly.

Both of these deficiencies are addressed in this paper. A lower bound on $d_1(\mu^*)$ is established in section 7, which overcomes the second of these shortcomings. With regard to the first, while we can prove a quantitative form of the shortest path characterization theorem, the associated path complexity bound turns out to be poorly suited for the purposes of constructing an efficient ε-approximation algorithm because of its dependence on L. (Recall that the characterization of shortest paths for a point or disc has no such dependence.) The development of an ε-approximation algorithm for 1ORM that depends only on a qualitative shortest path characterization theorem was the driving motivation for pseudo approximation framework that forms the key innovation of this paper.

This general framework shows how to start with any pseudo approximation algorithm satisying some simple properties and derive an efficient ε-approximation algorithm. While the formal definition of "pseudo approximations" is somewhat specialized for the current applications, it should be understood that the intuitive concept of a pseudo approximation is that of having an arbitrary additive term which can be made as small as one likes. In this sense, almost any straightforward discretization of a continuous optimization problem is a pseudo approximation. Therein lies the power of this framework.

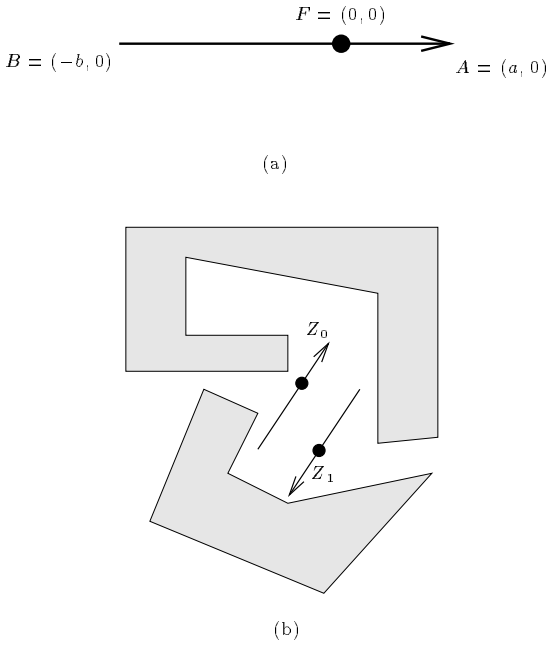


Figure 1: (a) Canonical position of rod AB with focus F. (b) Moving from Z_0 to Z_1 amidst polygonal obstacles (shaded area).

Optimum Motion Planning of a Rod. The configuration space for the motion of a rod is 3-dimensional (as for 3ESP) but it is non-Euclidean. A rod is a fixed directed line segment AB of unit length, as shown in Figure 1(a). The problem is to find a motion of the rod from some initial placement Z_0 to some final placement Z_1 while confined to a closed polygonal region $\Omega \subseteq \mathbb{R}^2$. The complementary set $\mathbb{R}^2 \setminus \Omega$ is the obstacle set. See Figure 1(b). Let F be a fixed point on AB, called the **focus** point. If μ is a motion of the rod, the corresponding trajectory of F is called the **trace** of μ; the length of this trace is the d_1**-distance** of μ, denoted by $d_1(\mu)$. If μ^* has the minimum d_1-distance among motions from Z_0 to Z_1, we write $d_1(Z_0, Z_1) = d_1(\mu^*)$ and call μ^* a d_1-optimum motion. Thus the input to our algorithms is (Z_0, Z_1, Ω), plus other input parameters such as ε when appropriate. We refer to Appendix I for further background, including a description of the **free space** $FP = FP(\Omega)$. Although we focus on d_1-optimality, other notions of optimality have been considered in the interesting history of this problem:

- The oldest work here is **Kakeya's problem** [5], which asks for the smallest area swept by the rod while moving it from any position to its "dual position" (this is the position reached by rotating the rod 180° about its midpoint), in the absence of obstacles. The obvious rotation motion sweeps out an area of $\pi/4$ which turns out to be far from optimal.

- **Ulam's problem** [25, 15] is to minimize the average length of the trajectories of the two endpoints A, B while moving from Z_0 to Z_1. This is also in the absence of obstacles. Icking et al [13] introduced the Cauchy surface area formula as a tool for analyzing such motions. They define the d_n-distance for rod motion, where $n \geq 1$ is any integer or $n = \infty$. This distance is a metric for $n \geq 2$, and Ulam's problem corresponds to d_2.

- An optimality notion that is not based on distance is to maximize the the minimum clearance (i.e., distance to the nearest obstacle) of the motion. Here, O'Dunlaing et al [14, 22, 23] gave a quadratic time solution based on the retraction approach.

- Although the d_1-distance is not a metric, it is a natural and interesting measure. Restricted forms of d_1-motions were investigated by Papadimitriou and Silverberg [17], Sharir [21] and O'Rourke [15]. The surprising result of Asano et al [3] is that d_1-optimal motion is NP-hard whenever F lies in the relative interior of the rod. The authors recently succeeded in extending the NP-hardness result to the case when F is at an endpoint of the rod. This result will be presented in near future.

Contributions of this work.

(I) We introduce the framework of **pseudo approximation algorithms** and derive an efficient search scheme for converting any suitable pseudo ap-

proximation algorithm into a true ε-approximation algorithm. The approximation algorithm is precision sensitive.

(II) We construct an efficient ε-approximation algorithm for the d_1-optimal motion of a rod. This is based on a new and simplified strongly polynomial pseudo ε-approximation algorithm. Instead of Equation (3), our pseudo approximate motion μ satisfies

$$d_1(\mu) \leq d_1(\mu_R^*) + \varepsilon R \tag{4}$$

where $\varepsilon > 0$ and $R > 0$ are arbitrary parameters. Here μ_R^* is the optimal d_1-motion when the trace is restricted to a ball of radius R centered at the initial position of the focus (denoted by $F[Z_0]$).

(III) As another application, we provide a new ε-approximation algorithm for the shortest path for a point amidst polyhedral obstacles in 3D. This algorithm depends logarithmically on the input precision parameter L. Previous ε-approximation algorithms are polynomial (or even exponential) in L. Being simpler than previous solutions, it is possibly implementable (a hope that was expressed for the algorithm in [16]). See also [20] for some experimental results.

(IV) An improved analysis of the boundary ∂FP of the space of free placements of a rod. This boundary is important for d_1-optimal motion. We introduce a 2-complex structure for this set.

Paper Overview. In Section 2, we introduce the framework of pseudo approximations. Section 3 shows how to convert a pseudo approximation algorithm into an ε-approximation algorithm. Section 4 gives a pseudo approximation algorithm for 3ESP, leading to a new ε-approximation algorithm. Section 5 reviews known results about the local structure of d_1-optimal rod motion. Based on this, we present in Section 6 a pseudo approximation algorithm for d_1-optimal rod motion (1ORM). Section 7 proves a lower bound on any non-zero d_1-distance. We conclude in Section 8. Appendices I and II summarize some background on d_1-optimal rod motion.

2 The Pseudo Approximation Framework

We describe an abstract framework for pseudo approximation algorithms.

Suppose that we are searching for an optimum solution x^* in a search space X, where optimality is based on minimizing the function $d : X \to \mathbb{R}_{\geq 0}$. Assume that the space X that has been "stratified" into the sets $\{X_R \subseteq X : R \geq 0\}$ such that (a) $X_R \subseteq X_{R'}$ for $R < R'$, and (b) we know a value R^* such that $X = X_{R^*}$. The parameter R should be interpreted as "search radius".

For each $R \geq 0$, let $x_R^* \in X_R$ denote an optimum solution in X_R. It follows that

$$d(x_R^*) \geq d(x_{R'}^*) \geq d(x^*) \tag{5}$$

for $R < R'$. Now assume that we have a **pseudo approximation function**,

$$\pi : (0,1] \times \mathbb{R}_{\geq 0} \to X$$

such that for all $\varepsilon \in (0,1]$ and $R \geq 0$, $\pi(\varepsilon, R) \in X_R$ (and hence $d(\pi(\varepsilon, R)) \geq d(x^*_R)$) and

$$d(\pi(\varepsilon, R)) \leq d(x^*_R) + \varepsilon R. \tag{6}$$

Since the parameter $\varepsilon > 0$ in (6) is fixed throughout this discussion, we may suppress it: we simple write "$\pi(R)$" instead of "$\pi(\varepsilon, R)$" where we view π as the function $\pi : \mathbb{R}_{\geq 0} \to X$.

There is one additional property we need:

$$d(x^*) \leq R \quad \Longrightarrow \quad d(x^*_R) = d(x^*). \tag{7}$$

This says that the search radius parameter R has some direct correlation with $d(x^*_R)$ (i.e., the stratification of X is tied to the underlying cost function).

Some Claims. To interpret what the following claims say, it is convenient to call R a **high value** in case $d(\pi(R)) \leq R$, and a **low value** otherwise.

Claim 1 *If R is a high value, i.e., $d(\pi(R)) \leq R$, then $R \geq d(x^*)$ and $d(x^*) = d(x^*_R)$.*

Proof. The first relation follows from $R \geq d(\pi(R)) \geq d(x^*_R) \geq d(x^*)$. The second relation $d(x^*_R) = d(x^*)$ is then a consequence of the "correlation property" (7). **Q.E.D.**

Claim 2 *If R is a low value, i.e., $d(\pi(R)) > R$, then $d(x^*) > R(1-\varepsilon)$.*

Proof. By way of contradiction, suppose $d(x^*) \leq R(1-\varepsilon)$. Then by (7), $d(x^*_R) = d(x^*)$. Hence $d(\pi(R)) \leq d(x^*_R) + \varepsilon R = d(x^*) + \varepsilon R \leq R(1-\varepsilon) + \varepsilon R = R$. This gives the desired contradiction. **Q.E.D.**

Claim 3 *Fix any constant $\alpha > 1$. If $R_{\min}$ is a low value, $R_{\max}$ a high value and $R_{\max} \leq \alpha R_{\min}$, then $d(\pi(R_{\max})) < d(x^*)\left(1 + \alpha \frac{\varepsilon}{1-\varepsilon}\right)$.*

Proof. Since $R_{\min}$ is a low value it follows from Claim 2 that

$$d(x^*) > (1-\varepsilon) R_{\min}. \tag{8}$$

Thus

$$\begin{aligned}
d(\pi(R_{\max})) &\leq d(x^*_{R_{\max}}) + \varepsilon R_{\max} && \text{(Equation (6))} \\
&= d(x^*) + \varepsilon R_{\max} && \text{(Claim 1)} \\
&\leq d(x^*) + \varepsilon \alpha R_{\min} \\
&< d(x^*) + \varepsilon \tfrac{\alpha\, d(x^*)}{1-\varepsilon} && \text{(Equation (8))} \\
&= d(x^*)\left(1 + \tfrac{\alpha\,\varepsilon}{1-\varepsilon}\right).
\end{aligned}$$

Q.E.D.

COROLLARY 1 *Assuming $\varepsilon \leq 1/2$, if $R_{\min}$ is a low value, $R_{\max}$ a high value and $R_{\max} \leq 2R_{\min}$, then*

$$d(\pi(R_{\max})) < d(x^*)(1+3\varepsilon) \tag{9}$$

We will also need the following:

CLAIM 4 *If R is a low value and $d(x^*_R) = d(x^*)$, then $d(\pi(R)) < d(x^*)/(1-\varepsilon)$.*

Proof. By Claim 2, $d(\pi(R)) > R$ implies $d(x^*) > R(1-\varepsilon)$. Thus $d(\pi(R)) \leq d(x^*_R) + \varepsilon R = d(x^*) + \varepsilon R < d(x^*) + \varepsilon d(x^*)/(1-\varepsilon) = d(x^*)/(1-\varepsilon)$. **Q.E.D.**

Suppose our pseudo-approximation function $\pi(R)$ has the following monotonicity property:

$$d(\pi(R-r)) \geq d(\pi(R)), \quad \text{for all } r \geq 0. \tag{10}$$

This property is not hard to ensure in the two main examples in this paper. It has an interesting consequence:

CLAIM 5 *If the pseudo approximation function satisfies (10), then it induces a $0-1$ ordering of the real numbers in the following sense: if R is a low value then $R-r$ is a low value, for all $r \geq 0$.*

Proof. If R is a low value, then $d(\pi(R)) > R$. Hence $d(\pi(R-r)) \geq d(\pi(R)) > R > R-r$, i.e., $R-r$ is a low value. **Q.E.D.**

3 Conversion to Precision Sensitive ε-Approximations

We now show how any pseudo approximation algorithm in the previous section can be converted into an ε-approximation algorithm.

Our analysis uses two simple assumptions: (1) $d(\pi(0)) > 0$ (i.e. 0 is a low value) and (2) $\varepsilon \leq 1/2$. Note that if (1) fails, then $d(\pi(0)) = 0$ and $\pi(0)$ is already an optimal solution. As for (2), we could have used any $C < 1$ in place of $1/2$.

We first present a simple binary search method. The idea is to exploit the Corollary to Claim 3, by maintaining a pair $(R_{\min}, R_{\max})$ satisfying the *invariant* that $R_{\min}$ is a low value and $R_{\max}$ a high value. We start from the largest search radius R^*. At each iteration, we halve the gap $R_{\max} - R_{\min}$, halting when the gap is at most $R_{\min}$. The basic comparison of the binary search is testing if a number $R > 0$ is high or low. This amounts to a call of the pseudo approximation function to compute $d(\pi(R))$ and comparing it with R.

SIMPLE BINARY SEARCH
OUTPUT: R such that $d(\pi(R)) \leq d(x^*)(1+3\varepsilon)$.
BASE CASE: if R^* is low, Return R^*.
$R_{\min} \leftarrow 0$ and $R_{\max} \leftarrow R^*$.
Do {
$R \leftarrow (R_{\max} + R_{\min})/2$.
If R is low, then $R_{\min} \leftarrow R$
else $R_{\max} \leftarrow R$.
} while $(R_{\max} > 2R_{\min})$
Return $R_{\max}$.

Correctness: The base case is justified by Claim 4 (use the fact $1/(1-\varepsilon) = 1+\varepsilon+\varepsilon^2/(1-\varepsilon) \leq 1+2\varepsilon$). The loop invariant is maintained as we update $R_{\min}$ and $R_{\max}$. Upon termination, the output is correct by the corollary of Claim 3.

Complexity: We claim that the number of iterations is at most $2+\lg(R^*/d(x^*))$. Each iteration reduces the gap $R_{\max} - R_{\min}$ by a factor of two, and the initial gap is R^*. So it suffices to show that the final gap is at least $d(x^*)/4$. Let $r_{\max}$ and $r_{\min}$ be the values of the variables $R_{\max}$ and $R_{\min}$ in the previous iteration (i.e., the last iteration for which $R_{\max} > 2R_{\min}$). Now, $(r_{\max} + r_{\min})/2$ is equal to either $R_{\max}$ or $R_{\min}$. Hence, either $R_{\max} = (r_{\max} + r_{\min})/2$ and $R_{\min} = r_{\min}$ or $R_{\max} = r_{\max}$ and $R_{\min} = (r_{\max} + r_{\min})/2$. Therefore, since $r_{\max} > 2r_{\min}$,

$$R_{\max} - R_{\min} = (r_{\max} - r_{\min})/2 > r_{\max}/4 \geq R_{\max}/4 \geq d(x^*)/4$$

(the last inequality follows from Claim 1). This concludes the proof.

Geometric Search. We can significantly speed up the above search using the following 2-tiered search:

GEOMETRIC SEARCH

OUTPUT: R such that $d(\pi(R)) \leq d(x^*)(1+3\varepsilon)$.

0. BASE CASE: if (R^* is low) then Return R^*.
1. KEY TEST: If ($R = 1$ is low)
   then $R_{\min} = 1$, $k = \lceil \lg \lg R^* \rceil$, and go to Search 2 directly.
   else $R_{\min} = 1/2$, $k = 0$, and go to Search 1 first.
2. SEARCH 1: // Now $R_{\min} 2^{2^k}$ is high
   While ($R_{\min}$ is high) do
   $R_{\min} \leftarrow R_{\min} 2^{-2^k}$; $k \leftarrow k+1$;
   // Invariant 1: $R_{\min} 2^{2^k}$ is high
3. SEARCH 2: // Now $R_{\min}$ is low and $R_{\min} 2^{2^k}$ is high
   $lo \leftarrow 0$; $hi \leftarrow 2^k$;
   While ($hi - lo > 1$) do
   // Invariant 2: $R_{\min} 2^{lo}$ is low, $R_{\min} 2^{hi}$ is high
   $m \leftarrow (hi + lo)/2$;
   if ($R_{\min} 2^m$ is low) then $lo \leftarrow m$;
   else $hi \leftarrow m$;
   $R_{\min} \leftarrow R_{\min} 2^{lo}$; $R_{\max} \leftarrow R_{\min} 2^{hi}$;
4. Return $R_{\max}$.

Correctness: The base case is justified as for the simple binary search. The two loop invariants are easily verified. The final return statement is again justified by the Corollary to Claim 3.

Complexity: After $k \geq 1$ iterations in Search 1, the variable $R_{\min}$ reaches the value $r_k := 2^{-2^0} 2^{-2^0} 2^{-2^1} \cdots 2^{-2^{k-1}} = 2^{-2^k}$. If r_k is high, then $r_k \geq d(x^*)$; hence $2^{2^k} \leq 1/d(x^*)$ and so $k \leq \lg \lg(1/d(x^*))$. But the total number of iterations in Search 1 is just one more than the last value of k for which r_k is high. This proves the first search has $\leq 1 + \lg \lg(1/d(x^*))$ iterations. In Search 2, we begin with a pair of values $(lo, hi) = (0, 2^k)$. If we came directly from Step 1 (Key Test), then $k = \lceil \lg \lg R^* \rceil$; otherwise we just completed Search 1 and $k \leq \lfloor \lg \lg(1/d(x^*)) \rfloor$. We conclude that

$$k \leq \max\{\lceil \lg \lg R^* \rceil, \lfloor \lg \lg(1/d(x^*)) \rfloor\}.$$

Initially $hi - lo = 2^k$ but finally we have $1/2 < hi - lo \leq 1$. The number of iterations is $\leq 1 + k = 1 + \max\{\lceil \lg \lg R^* \rceil, \lfloor \lg \lg(1/d(x^*)) \rfloor\}$.

In summary:

LEMMA 2 *The geometric search procedure determines R such that $d(\pi(R)) \leq d(x^*)(1+3\varepsilon)$ using at most*

$$2 + \max\{\lceil \lg \lg R^* \rceil, 2 \lfloor \lg \lg(1/d(x^*)) \rfloor\}$$

calls to the pseudo approximation function.

Let us simply write $\max\{\lceil \lg \lg R^* \rceil, 2\lfloor \lg \lg(1/d(x^*)) \rfloor = \Theta(\lg \lg(R^*/d(x^*)))$ As corollary, if computing $\pi(R)$ and $d(\pi(R))$ takes time $T = T(\varepsilon, R^*)$, we achieve an ε-approximation algorithm whose time is $O(T \times \lg \lg(R^*/d(x^*)))$.

Note that Search 2 can be viewed as a "geometric mean" search: we can view the search as maintaining an interval $(R_{\min}, R_{\max})$ in which each "halving step" involves replacing one endpoint of the interval by the geometric mean $\sqrt{R_{\min} R_{\max}}$. Even Search 1 can be interpreted in this way (except that in the absense of a low value, we replace $R_{\min}$ by 1). This explains why we call this search method "Geometric Search".

Note that when R is high, the combination of Search 1 and Search 2 is really a disguised form of "unbounded search" in the sense of Bentley and Yao [4]. The difference is that standard unbounded search uses an absolute error bound, while we use a relative error bound (cf. Corollary to Claim 2). Our complexity bound is $2 + \lfloor \lg \lg(1/d(x^*)) \rfloor$. Known techniques for unbounded searching can be applied here to improve the upper bound to $\lfloor \lg \lg(1/d(x^*)) \rfloor + o(\lg \lg(1/d(x^*))$. We leave such improvements as an exercise.

One variant is to compute an *á priori* lower bound r^* on $d(x^*)$. We can perform a simultaneous search for critical value radius from above (starting from R^*, as in Search 1) *and* from below (starting from r^*). The number of iterations is then

$$\min\{\lg(\lg(R^*/d(x^*)), \lg \lg(d(x^*)/r^*)\}.$$

In this way, we lose precision sensitivity but gain the potential to have speed up when $d(x^*)$ is very large (near R^*) or very small (near r^*).

Precision-Sensitive Solution. The running time of our ε-approximation algorithm depends on $d(x^*)$. In our applications below, $R^* \leq c^L$ for some $c > 1$ and the maximum bit length L of the input numbers. Hence if $d(x^*) < c^{-L}$ then $\lg \lg(R^*/d(x^*)) = \Theta(\lg \lg(1/d(x^*)))$; otherwise $\lg \lg(R^*/d(x^*)) = \Theta(\lg \lg(R^*))$. We call $\lg(1/d(x^*))$ the "output precision" of the problem instance since it is proportional to the number of bits needed to express $d(x^*)$ to within a constant factor. In this sense, we say that our algorithm is **precision sensitive**. Note that despite being sensitive to $d(x^*)$, our algorithm does not explicitly know x^* or $d(x^*)$.

The notion of precision sensitivity was first introduced in Choi et al [8]. It should be noted that the 3ESP approximation algorithm there has a stronger objective than our version here: their objective was to determine a sequence of edges that determines the shortest path, as well as to determine an ε-approximate feasible motion along this sequence. Call this the "Combinatorial 3ESP Approximation Problem". In this version, it is natural to define the "output precision" as $\log(1/\Delta)$ where $\Delta = d(x^*) - d(x_2)$ and x_2 is a combinatorially distinct next shortest path. The gap Δ is a measure of the necessary number of bits that must be evaluated if we want to distinguish x^* from x_2. The current best lower bound for non-zero Δ is doubly exponentially small (so $\lg(1/\Delta)$ is single exponential). The precision-sensitive algorithm of [8] is only polynomial in the output precision, and hence the overall algorithm is not known to be

polynomial time. In contrast, for our two applications below, we show a linear upper bound on $\lg(1/d(x^*))$ and so the running time of our algorithms are actually logarithmic in the output precision.

4 Approximation Algorithm for 3ESP

We now construct a pseudo approximation algorithm for 3ESP that has the properties required for our binary search method in the previous section. This serves three purposes: (1) It gives the first indication that our abstract setting in the previous section is non-vacuous. (2) It will serve as a model for our next section, where we give a pseudo approximation algorithm for rod motion. Additional complications will arise in the case of rod motion. (3) Finally, our new algorithm has advantages over previous solutions (some of which address shortest path problems in more general settings) [16, 9, 8, 1, 2, 20, 24]. One is its simplicity, thus making it a more likely candidate for implementation. Another is its running time complexity being $O(\log L)$, in contrast to previous algorithms whose complexity is polynomial (or even exponential) in L.

The main result of this section is this:

THEOREM 3 *There is a pseudo approximation algorithm for 3ESP whose running time is $O(n^4/\varepsilon^2)$, or more precisely $O(n^4\varepsilon^{-2}\lg\lg(2^L/d(x^*)))$.*

This is a worst case complexity bound: the actual complexity is $O(\min\{n^2/\varepsilon,\ |A|\log(n^2/\varepsilon)\})$, where $|A|$ denotes the number of edges in the graph that is searched to find the shortest path. It is also worth noting that this complexity does not depend on the parameter R (at least in the algebraic complexity model). While it is possible to convert this bound into the bit complexity model, a tight analysis can be intricate. Reference [8] is one of the few papers that gave a careful accounting of the bit complexity of an approximation algorithm.

Coupled with the geometric search procedure in the previous section, and using the estimate $R^* = 2^L$, we obtain an ε-approximation algorithm with complexity

$$O(n^4\varepsilon^{-2}\lg\lg(2^L/d(x^*))).$$

As this expression depends on $d(x^*)$, to obtain an *á priori* bound on the running time, we need a lower bound on $d(x^*)$ in terms of the input parameters. This is not hard: we may assume $Z_0 \neq Z_1$ or else the problem is trivial. Then we obtain $d(x^*) \geq \|Z_0 - Z_1\| \geq c^{-L}$, for some constant $c > 1$.

Input Parameters and Representation. Let $R \geq 1$ and $\varepsilon' > 0$ be given. (Later we will choose $\varepsilon' = \varepsilon/6$ where ε is the corresponding pseudo approximation parameter.) There is also the the standard input (Z_0, Z_1, Ω) for motion planning, where $Z_0, Z_1 \in \mathbb{R}^3$ and $\Omega \subseteq \mathbb{R}^3$ is a closed polyhedral set. There are n obstacle edges and vertices, and the numbers used in their description are L-bit rational numbers, *i.e.*, the numerators and denominators are L-bit integers. We shall write Ω_R for the restriction of Ω to a ball of radius R centered at Z_0. Let μ_R^* denote a shortest path from Z_0 to Z_1 when restricted

to Ω_R. Our goal is to compute a path in Ω_R that is a "pseudo approximation" to μ_R^*.

Fragment Visibility Graph. We describe a fragmentation of the obstacle edges (that depends on the parameters R, ε' and n) and a "visibilty graph" $FVG = (N, A; W)$ on the resulting set of fragments. Specifically, FVG is a weighted undirected graph, N is the **node set**, A the **arc set**, and $W : A \to \mathbb{R}_{>0}$ is the weight function. Note that we use the node/arc terminology to avoid conflict with the vertex/edge terminology reserved for the obstacle set Ω. First, we specify the node set N. For each obstacle edge $e \in \Omega$, we replace it with $e_R = e \cap \Omega_R$. Then we subdivide e_R into $\lceil n/\varepsilon' \rceil$ **fragments**, each of length at most $R\varepsilon'/n$. Summed over all edges, we have $O(n^2/\varepsilon')$ fragments. Next, the node set N comprises these fragments, including Z_0 and Z_1 as special fragments. The arc set A comprises those pairs (σ_a, σ_b) of fragments that are "weakly visible", i.e., there is a line segment $[\sigma_a^*, \sigma_b^*] \subseteq \Omega_R$ such that $\sigma_a^* \in \sigma_a$ and $\sigma_b^* \in \sigma_b$. We define the arc's weight $W(\sigma_a, \sigma_b)$, not as $\|\sigma_a^* - \sigma_b^*\|$, but as $\|\check{\sigma}_a - \check{\sigma}_b\|$ where $\check{\sigma}_a$ and $\check{\sigma}_b$ are the midpoints of σ_a and σ_b. We refer to [8] for the details of these computations (deciding weak visibility, etc).

Given the fragment visibility graph $FVG = (N, A; W)$, we can compute the shortest path P^* from the source Z_0 to Z_1 using standard techniques such as Dijkstra's algorithm. Since $|A| = O(n^4/\varepsilon'^2)$ this algorithm can be implemented in time $O(|N| \log |N| + |A|) = O(n^4/\varepsilon'^2)$, using Fibonacci heaps (E.g., [10]).

Relation between the Fragment Visibility Graph shortest path P^* and the restricted Euclidean shortest path μ_R^*. The relation is indirect: from P^*, we construct an Euclidean motion $\mu(P^*)$. Then we derive a relation between $\mu(P^*)$ and the restricted Euclidean shortest path μ_R^*. First we set out some basic properties that follow immediately from the definition of the fragmentation process:

Observation A. If P is any single edge path in FVG joining two fragments σ_a and σ_b of different obstacle edges, then their exists an Euclidean path $\mu(P)$ in Ω_R joining $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + \varepsilon' R/n$.

Observation B. If P is any path in FVG joining two fragments σ_a and σ_b of the same obstacle edge, then their exists an Euclidean path $\mu(P)$ in Ω_R joining $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + \varepsilon' R/n$.

Observation C. If Euclidean line segment μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b of different obstacle edges, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + \varepsilon' R/n$.

Observation D. If Euclidean path μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b of the same obstacle edge, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + \varepsilon' R/n$.

Let P be any path in the FVG from Z_0 to Z_1. We can decompose P into $k \leq 2n$ subpaths $P_1, P_2, \ldots, P_k$, such that odd-indexed subpaths consist of single edges joining fragments of different obstacle edges, and even-indexed subpaths (possibly of zero length) join fragments of the same obstacle edge. Define $\mu(P)$

to be the Euclidean path formed by concatenation of $\mu(P_1), \ldots, \mu(P_k)$. By Observations A and B, we have:

$$\begin{aligned} d(\mu(P) &= \sum_{i=1}^{k} d(\mu(P_i)) \\ &\leq \sum_{i=1}^{k} [W(P_i) + \varepsilon' R/n] \\ &\leq W(P) + 2\varepsilon' R. \end{aligned} \tag{11}$$

On the other hand, if μ is any Euclidean path in Ω_R then μ can be decomposed into $k \leq 2n$ subpaths $\mu_1, \mu_2, \ldots, \mu_k$ such that odd-indexed subpaths consist of single line segments joining fragments of different obstacle edges, and even-indexed subpaths (possibly of zero length) join points on fragments of the same obstacle edge. Define $P(\mu)$ to be the path in FVG formed by concatenation of the paths $P(\mu_1), \ldots, P(\mu_k)$. By Observations C and D, we have:

$$\begin{aligned} W(P(\mu)) &= \sum_{i=1}^{k} W(P(\mu_i)) \\ &\leq \sum_{i=1}^{k} [d(\mu_i) + \varepsilon' R/n] \\ &\leq d(\mu) + 2\varepsilon' R. \end{aligned} \tag{12}$$

We now put these two steps together. Begin with the shortest path P^* joining Z_0 and Z_1 in FVG. Using the construction above, we obtain an Euclidean path $\mu(P^*)$.

LEMMA 4

$$d(\mu(P^*)) \leq d(\mu_R^*) + 4\varepsilon' R.$$

Proof.

$$\begin{array}{rcll} d(\mu(P^*)) & \leq & W(P^*) + 2\varepsilon' R & \text{(by (11))} \\ & \leq & W(P(\mu_R^*)) + 2\varepsilon' R & \text{(by definition of } P^*) \\ & \leq & d(\mu_R^*) + 4\varepsilon' R & \text{(by (12))} \end{array}$$

Q.E.D.

We make the final connection to the abstract binary search of the previous section. The restricted search space X_R comprise all the Euclidean paths in Ω_R. The pseudo approximation function $\pi : (0, 1] \times \mathbb{R} \to X$ is given by $\pi(\varepsilon, R) = \mu(P^*)$ where P^* is the shortest path in the Fragment Visibility Graph constructed with the parameters $\varepsilon' = \varepsilon/4$ and R. By lemma 4, we know that π is a pseudo approximation function. We need to verify the properties $d(\mu^*) \leq R$ implies $d(\mu_R^*) = d(\mu^*)$ (Equation (7)) and $R^* \leq 2^L$. But they are easily seen. This proves the main result stated above.

REMARK: This approach immediately generalizes to shortest path for a point robot moving amidst polyhedral obstacles in any fixed dimension.

5 Local Structure of Optimum Rod Motion

In this section, we give a top level overview of the underlying geometry of optimal d_1-motion of a rod. This is necessary prerequisite for the approximation algorithm to be presented in the next section. We defer details to the appendices because of some overlap with [3]. However, the current paper offers technical improvements and alternative treatment.

Decomposition of ∂FP. There are three known approaches to analyzing the structure of the free space $FP = FP(\Omega)$ of a rod. In [18], the set FP is decomposed into **cells** using the notion of concept of clockwise- and counterclockwise-stops. The projection of these cells on to $\mathbb{R}^2$ gives a simple decomposition of Ω into planar regions. The second approach [22, 23] is based on the concept of Voronoi cells. Neither of these approaches are suitable for analyzing optimum d_1-motion. The third approach [3] was introduced for d_1-optimal motion, and is based on a cell decomposition of ∂FP. This boundary can be decomposed into a 2-complex comprising 2-cells (called **patches**), 1-cells (called **edges**) and 0-cells (called **vertices**). The details are given Appendix II. In particular, it is shown (theorem 8) that the number of patches, edges and vertices in the complex is $O(n^2)$, $O(n^3)$ and $O(n^3)$, respectively.

Similar to retraction-based motion planning [14], our goal is to construct a 1-dimensional complex (i.e., a skeleton) in which optimum motion can be found. The obvious place to look is to consider the subcomplex K of ∂FP comprising the edges and vertices. The skeleton we seek will clearly need to augment K with additional edges to ensure global connectivity. But it turns out that no finite number of edges suffices for d_1-optimum motion.

On Stopover Curves and Mirrors. One can still hope for some "parametrized form" of the skeleton that can reduce optimal motion to a finite graph search. To see how this might work, consider the well-understood problem of optimum motion μ^* for the unit disc in the plane [12]. Here, the circular arcs of radius $1/2$ and centered at convex corners are called **displaced corners** [3]. They are important because the trace of μ^* is non-straight only by incorporating parts of these arcs. The optimal trace will join and leave such circular arcs at a tangent. If every tangent is potentially part of an optimum motion, there is no finite graph that we can construct. But it turns out that we can construct a finite search graph in which these displaced corners are used as "parametrized vertices". The analog of these displaced corners for optimal d_1-optimal rod motion is called **stopover curves** [3]. Each stopover curve is simply the trace of a corresponding **stopover edge** of ∂FP. Each placement Z in a stopover curve is **constricted** in the sense that rotation about the focus F is impossible: the rod is stopped by two obstacle features, one acting as a clockwise stop, the other as a counter clockwise stop. There are three kinds of stopover curves: (1) **circular edge** defined by a concave corner C, (2) **elliptic edge** defined by a pair to distinct walls W_1, W_2, and (3) **conchoidal edge** defined by a convex corner C and a wall W. The traces of these edges are parts of a circle, an ellipse

or an conchoid. In addition to these cases, there is a fourth kind of constricted edge: these are defined by two corners and the trace is a straight line segment. We do not consider this a stopover curve. See Appendix II for more details.

The number of stopover edges is $O(n^3)$ and we might hope for a finite search graph using these curves as "parametrized vertices" in the search graph. But this hope is dashed by the phenomenon of **mirrors**. In a d_1-optimal motion, the traces can "reflect" off the mirrors following the law of reflection (c.f. Appendix II). In fact, mirrors are the reason for NP-hardness [3]. Fortunately, when we seek approximate motions, we can avoid mirrors altogether.

Locally d_1-optimal motion. In order to make further progress, we need to understand what can happen "locally" in a d_1-optimal motion $\mu : [0,1] \to FP$. Intuitively, the trace $F\mu : [0,1] \to \mathbb{R}^2$ must travel along a straight line unless it is forced by some obstacle features to turn or bend.

Let X be a metric space with metric $d(x,y)$, and fix a continuous function (curve) $f : [0,1] \to X$. If $\varepsilon > 0$ and $0 < t_0 < 1$, we call an open interval I an (ε, f)**-neighborhood** of t_0 if $t_0 \in I$ and for all $t \in I$, $d(f(t), f(t_0)) < \varepsilon$. We may also call I an f**-neighborhood** of t_0 if it is a (ε, f)-neighborhood for some $\varepsilon > 0$. If P is any property of curves, we say that f **satisfies** P **at** t_0 $(0 < t_0 < 1)$ if there exists $\varepsilon = \varepsilon(t_0) > 0$ and a (ε, f)-neighborhood I of t_0 such that the restriction $f|I$ of f to I satisfies property P. If f satisfies P at all t_0 $(0 < t_0 < 1)$, then we say f **satisfies** P **locally**.

An interval $I \subseteq [0,1]$ is said to be **stationary** for f if $f|I$ is a constant function. Our definition of neighborhood is necessitated by the presence of stationary intervals that are nontrivial (i.e., that have non-empty interior). The **essential f-neighborhood** of t_0 is the intersection of all f-neighborhoods of t_0. The essential f-neighborhoods of t_0 is stationary for f. In applications, we choose $X = \mathbb{R}^2$ with the Euclidean metric, and $f = F\mu$ where μ is any motion. Often, we want to focus on the properties of μ, not $F\mu$. To do this, we use $F\mu$-intervals while discussing properties of μ, as seen next. We define three local properties of motions:

- μ is Locally Straight: Let $P_0(\mu)$ be the property that "the trace $F\mu$ is straight". We say a motion μ is **locally straight at** t if P_0 holds whenever F is restricted to a $F\mu$-neighborhood of t.

- μ is Locally a Vertex: Let $P_1(\mu)$ be the property "there exists $0 \leq t \leq 1$ such that $\mu(t)$ is a vertex (0-cell)". Then we say μ is **locally a vertex at** t if P_1 holds whenever μ is restricted to a $F\mu$-neighborhood of t. Equivalently, μ is locally a vertex at t iff $\mu(t)$ is a vertex for some t in the $F\mu$-essential neighborhood of t.

- μ is Locally Reflecting: Let $P_2(\mu)$ be the property that "there exists $0 \leq t \leq 1$ such that $\mu(t)$ is reflecting". We say that μ is **locally reflecting at** t if P_2 holds whenever μ is restricted to a $F\mu$-neighborhood of t.

Here is the statement of the main result:

THEOREM 5 (LOCAL CHARACTERIZATION) *Let $\mu : [0,1] \to FP$ be a d_1-optimal motion and $0 < t_0 < 1$. Suppose μ is* not *locally straight at t_0. Then one of the following four situations hold:*

1. *μ is locally a vertex at t_0.*
2. *$\mu(t_0)$ is pivotal at a convex corner C (i.e., $F[\mu(t_0)] = C$). Moreover, $F\mu$ is locally "bending" around C. See Figure 2(a).*
3. *$\mu(t_0)$ is constricted. Either (a) $F\mu$ is is locally tracing a stopover curve at t_0, or (b) the trace is a straight line that meets a stopover curve tangentially at t_0. See Figure 2(b).*
4. *μ is locally reflecting at t_0. See Figure 2(c,d). Let the curve γ be the displaced wall or corner where this reflection takes place. Then the trace $F\mu$ meets and leaves γ at an incident point r according to Snell's law. Morover, the trace in the neighborhood of r lies within the zone of γ.*

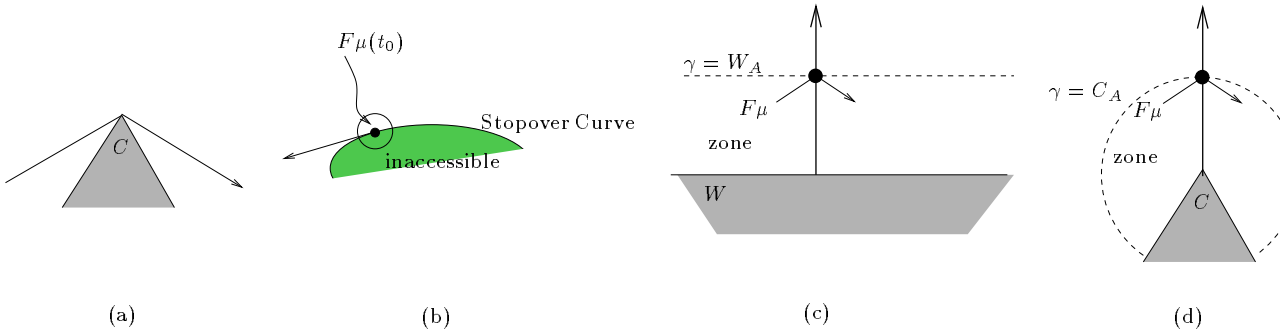


Figure 2: Locally non-straight traces: (a) pivoting around corner, (b) tangential or tracing a stopover curve, (c,d) reflecting off a displaced wall or corner

Note that this theorem does not say anything about the motion when μ is locally straight at t_0. In [3] we stated such a result without proof. We now provide the proof in Appendix III.

6 Pseudo Approximation Algorithm for Rod Motion

The main result of this section is the following.

THEOREM 6 *There is a pseudo approximation algorithm for d_1-optimal rod motion whose running time is $O(n^5\varepsilon^{-2}(n+\varepsilon^{-1}))$.*

Its proof follows the model of the 3ESP algorithm. We will define a fragmentation (depending on the input parameters (Z_0, Z_1, Ω), $\varepsilon > 0$ and $R \geq 0$) of the edges that bound the space of free configurations, and a "visibility graph" $FVG = (N, A; W)$ on these fragments. Specifically, if $FP = FP(\Omega)$ is the space of free configurations, define the **restricted free space** to be $FP_R = \{Z \in FP : F[Z] \in \Omega_R\}$ where Ω_R is defined as before, namely $\Omega_R = \Omega \cap B(R)$ and $B(R)$ is the disc of radius R centered at $F[Z_0]$. A **restricted optimal motion** of a rod is a motion $\mu_R^* : [0,1] \to FP_R$ such that $\mu_R^*(i) = Z_i$ $(i = 0, 1)$ and $d_1(\mu_R^*)$ is minimum.

To specify the node set N, we consider edges e of the complex ∂FP that are either stopover curves or mirrors. Let e_R be the restriction of e to FP_R. We can view such an edge e as a motion $e : [0,1] \to FP$, and thus speak of the trace of e. We can subdivide e into submotions, again called **fragments**. The **length** of a fragment σ is simply the length of the trace $F\sigma$, denoted by $d_1(\sigma)$. The node set N comprising these fragments, together with pivotal edges and Z_0 and Z_1, has total size $O(n^{5/2}\varepsilon'^{-1}(n^{1/2} + \varepsilon'^{-1/2}))$. The fragmentation process depends on the type of the edge.

Straight Mirrors We let each fragment have length $\leq \varepsilon' R/n$. Each straight mirror is associated with a wall. Let W be a wall of FP_R. Then $|W| \leq R$. Suppose S_W comprise all mirror fragments associated with W. Then the total length of all the fragments in S_W is $|W|$, and thus $|S_W| \leq n/\varepsilon'$. Since there are $\leq n$ walls, the number of fragments from straight mirrors is $O(n^2/\varepsilon')$.

Stopover Edges Again we first break the edges into fragments of length $\leq \varepsilon' R/n$. Since every stopover edge has total length which is $O(1)$, and there are $O(n^2)$ stopover edges, the total number of stopover fragments is $O(n^3/\varepsilon')$.

Circular Mirrors This is more complicated. First, we create **superfragments** of length $\leq (\varepsilon'/n)^{1/2}$. Each superfragment is further subdivided into fragments of length $\leq (\varepsilon'/n)^{3/2}$. Each circular mirror is associated with a convex corner. Let C be such a corner and let S_C comprise all fragments associated with C. The length of all the fragments in S_C is $< 2\pi$, and hence $|S_C| = O((n/\varepsilon')^{3/2})$. Since there are $\leq n$ corners, the number of fragments from circular mirrors is $O(n^{5/2}\varepsilon'^{-3/2})$.

We define the arc set $A \subseteq N^2$ based on weak visibility again, but the low level computation is somewhat more complex (but $O(1)$ in the algebraic complexity model). As before the weight $W(\sigma_a, \sigma_b)$ of the arc joining weakly visible fragments σ_a and σ_b is taken to be the Euclidean distance between the fragment midpoints. In addition, there is an arc joining fragments from the same mirror superfragment or stopover edge whose length is just the distance along the mirror superfragment or stopover edge between the fragment midpoints. The connection between paths in the Fragment Visibility Graph and d_1-motion, through the analogues of (11), (12), is complicated by the non-linear features of our domain, in particular circular mirrors.

Observation A. If P is any single edge path in FVG joining two fragments σ_a and σ_b not both of which are circular mirror fragments, then their exists an Euclidean path $\mu(P)$ in Ω_R joining their midpoints $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + \varepsilon' R/n$.

Observation A'. If P is any single edge path in FVG joining two fragments σ_a and σ_b both of which are circular mirror fragments, then their exists an Euclidean path $\mu(P)$ in Ω_R joining their midpoints $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + (\varepsilon'/n)^{3/2} R$.

Observation B. If P is any path in FVG joining two fragments σ_a and σ_b of the same straight mirror or stopover edge, then their exists an Euclidean path $\mu(P)$ in Ω_R joining $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + \varepsilon' R/n$.

Observation B'. If P is any path in FVG joining two fragments σ_a and σ_b of the same circular mirror superfragment, then their exists an Euclidean path $\mu(P)$ in Ω_R joining $\check{\sigma}_a$ and $\check{\sigma}_b$ such that $d(\mu(P)) \leq W(P) + (\varepsilon'/n)^{3/2} R$.

Observation C. If Euclidean line segment μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b not both of which are circular mirror fragments, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + \varepsilon' R/n$.

Observation C'. If Euclidean line segment μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b both of which are circular mirror fragments, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + (\varepsilon'/n)^{3/2} R$.

Observation D. If Euclidean path μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b of the same straight mirror or stopover edge, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + \varepsilon' R/n$.

Observation D'. If Euclidean path μ in Ω_R joins two points v_a and v_b, belonging to fragments σ_a and σ_b of the same straight mirror superfragment, then there exists a path $P(\mu)$ in FVG joining fragments σ_a and σ_b such that $W(P(\mu)) \leq d(\mu) + (\varepsilon'/n)^{3/2} R$.

Let P be any path in the FVG from Z_0 to Z_1. We can decompose P into subpaths $P_1, P_2, \ldots, P_k$, where each P_i is either (i) a single edge joining fragments of distinct stopover edges, straight mirrors, or circular mirror superfragments, or (ii) maximal subpaths joining fragments of the same stopover edge, straight mirror, or circular mirror superfragment. By maximality each stopover edge, straight mirror, or circular mirror superfragment can be involved in at most two subpaths of type (i). Thus, there are $O(n)$ subpaths of type (i) involving at least one non circular mirror fragment, and $O(n(n/\varepsilon'))$ subpaths of type (i) joining circular mirror fragments. Similarly, there are $O(n)$ subpaths of type (ii) involving at least one non circular mirror fragment, and $O(n(n/\varepsilon'))$ subpaths of type (ii) joining circular mirror fragments. If we define $\mu(P)$ to be the Euclidean path formed by concatenation of $\mu(P_1), \ldots, \mu(P_k)$ then, by Observations A, A', B and B', we have

$$d(\mu(P) \quad \leq \quad W(P) + O(\varepsilon' R). \tag{13}$$

Similarly, if μ is any Euclidean path in Ω_R then μ can be decomposed into subpaths $\mu_1, \mu_2, \ldots, \mu_k$ such that each μ_i is either (i) a single line segments joining fragments of distict stopover edges, straight mirrors, or circular mirror superfragments, or (ii) a maximal subpath joining fragments of the same stopover edge, straight mirror, or circular mirror superfragment.

Define $P(\mu)$ to be the path in FVG formed by concatenation of the paths $P(\mu_1), \ldots, P(\mu_k)$. By Observations C, C', D and D', we have

$$W(P(\mu)) \quad \leq \quad d(\mu) + O(\varepsilon' R). \tag{14}$$

We now put these two steps together. Begin with the shortest path P^* joining Z_0 and Z_1 in FVG. Using the construction above, we obtain an Euclidean path $\mu(P^*)$. As before it follows directly from equations (13) and (14) that $d(\mu(P^*)) \leq d(\mu_R^*) + O(\varepsilon' R)$.

This concludes our proof of Theorem 6. Now we can apply the geometric search algorithm to obtain a true ϵ-approximation algorithm whose running time depends on $\lg(1/d_1(\mu^*))$. We next prove a lower bound on $d_1(\mu^*)$.

7 Lower Bound on non-zero $d_1(\mu^*)$

This section is devoted to the proof of the following result:

THEOREM 7 *Assume $d_1(\mu^*) > 0$. Then $d_1(\mu^*) \geq c^{-L}$ for some constant $c > 1$.*

First, recall that the input description, including all obstacle corners and the initial position $F[Z_0]$ of the rod focus point F, involves only L-bit rational numbers. We assume that the lengths of the half-rods, AF and FB, are also L-bit rational numbers, denoted a and b respectively.

The sparsity of points with small rational coordinates allows us to easily dispose of the case in which the initial and final positions of the focus differ: if $F[Z_0] \neq F[Z_1]$ then $d_1(Z_0, Z_1) \geq \|F[Z_0] - F[Z_1]\| \geq 2^{-2L}$. Henceforth, assume $F[Z_0] = F[Z_1]$. In this case the theorem is an immediate consequence of the following:

Main Claim. If $F[Z]$ is never too far (no further than c^{-L}, for some fixed constant $c > 1$) from $F[Z_0]$ then the placements $F[Z_0]$ and $F[Z_1]$ must be equivalent up to a pure rotation (i.e., $d_1(Z_0, Z_1) = 0$).

Imagine rotating the rod about its focus F. We refer to the first obstacle feature encountered in a clockwise (respectively, counterclockwise) rotation in configuration Z as the **CW** (respectively, **CCW**) **feature stop at** Z. Two configurations with coincident foci are equivalent up to a pure rotation if and only if their CW and CCW feature stops are identical. To establish this equivalence it suffices to consider each rod endpoint separately and show that the CW and CCW feature stops of each half rod are identical in the two configurations.

The argument will be made for the half rod AF; the corresponding argument for FB is identical. Start by considering an open disc D of radius a centered at $F[Z_0]$. Then every obstacle wall either (i) does not intersect D; (ii) intersects

the boundary of D twice (and so has no associated corner in D); (iii) intersects the boundary of D once (and so has exactly one corner in D); or (iv) is entirely contained (and so has both of its associated corners) in D. We will choose **representative points** associated with walls as follows: Walls of type (i) have no representative point. Walls w of type (ii) are represented by their closest point p_w to $F[Z_0]$, provided $p_w \neq F[Z_0]$. (If $p_w = F[Z_0]$ then w has no representative point.) Walls w of type (iii) are represented by their corner inside D, provided this corner does not coincide with $F[Z_0]$. (If the corner coincides with $F[Z_0]$ then w has no representative point.) Finally, walls w of type (iv) are represented by their two associated corners, provided the corner does not coincide with $F[Z_0]$.

Walls of type (ii) and (iii) that do not have representative points (because they pass through or end at $F[Z_0]$) serve to partition D into wedges centred at $F[Z_0]$. The wedge occupied by the rod in its initial configuration (the entire disc, if there are no edges of type (ii) or (iii)) is called the **primary wedge**. Clearly, unless $F[Z]$ exceeds distance a from $F[Z_0]$ somewhere in its trace, the rod also occupies the primary wedge in its final configuration. The representative points serve to partition the primary wedge of D into secondary wedges formed by rays centered at $F[Z]$ through the individual representative points. Note that as Z changes the secondary wedges change. However, if $F[Z]$ never leaves a disc of radius c^{-L}, where $c > 1$ is a fixed constant, then (i) the distance from $F[Z]$ to individual representative points never exceeds a, and (ii) the relative order of the rays defining the secondary wedges (except for those rays that coincide in the initial, and final, configuration) remains fixed. This follows from the following elementary observations:

Observation A. All representative points are described by $O(L)$-bit rational numbers.

Observation B. The distance from $F[Z_0]$ to its closest point on the radius a circle centered at any representative point is at least c^{-L}, for some positive constant $c > 1$.

Observation C. The distance from $F[Z_0]$ to its closest point on the line joining any two representative points is either zero or at least c^{-L}, for some positive constant $c > 1$.

It follows that the secondary wedges, Δ_0 and Δ_1, occupied by the half rod AF in its initial and final configurations are idential. It remains to show that any free placement Z of AF in Δ_0 with $F[Z] = F[Z_0]$ has a unique CW and CCW stop. Since Δ_0 contains no representative points in its interior, if the CW (or CCW) stop of AF at Z is realized by a obstacle corner then that corner must be a representative point defining one of the bounding rays of Δ_0. If the CW (or CCW) stop of AF at Z is realized by a wall w, we observe that: (i) since AF has a free placement in Δ_0, w does not intersect both sides of Δ_0; (ii) since Δ_0 contains no representative points in its interior, w does not have an associated corner inside Δ_0, or cross the boundary of D twice within Δ_0; and hence (iii) w must intersect the boundary of D and one of the sides of Δ_0. The uniqueness of w follows from the fact that obstacle walls do not intersect.

This concludes the proof of the Main Claim.

8 Final Remarks and Open Problems

This paper has demonstrated the usefulness of pseudo approximation algorithms. Applied to two of the simplest NP-hard optimum motion planning problems, we obtain new ε-approximation algorithms. Many interesting questions remain to be explored:

(1) There should be other application of our general methodology to exploiting pseudo approximations.

(2) Our complexity analysis is in the algebraic model of computation. It is of interest to obtain true bit-complexity bounds, in the spirit of Choi et al [8].

(3) The complexity of the two approximation algorithms (shortest path for a point robot in 3D and d_1-shortest path for a rod in 2D) should be possible to improve, by not treating the pseudo approximation algorithms as black-boxes in the binary search scheme.

(4) Can we extend these techniques to approximating general d_n-optimal motion, and also to an arbitrary rigid planar robot?

ACKNOWLEDGEMENT

The authors would like to thank John Hershberger for his many insightful comments on the paper. The Binary Searches of Section 2 are improved from the original proceedings, thanks to an insightful question from David Mount about the necessity of a $\log(1/\epsilon)$ term in our original complexity bound.

References

[1] L. Aleksandrov, M. Lanthier, A. Maheshwari, and J.-R. Sack. An ε-approximation scheme for weighted shortest paths. *Proc. 6th SWAT*, Lecture Notes in Computer Science 1432, Springer Verlag, 11–22, 1998.

[2] L. Aleksandrov, A. Maheshwari, and J.-R. Sack. Aapproximation scheme for geometric shortest path problems. *Proc. 32nd STOC*, 286–295, 2000.

[3] T. Asano, D. Kirkpatrick, and C. Yap. d_1-Optimal motion of a rod. *12th ACM Symposium on Computational Geometry*, pages 252–263, 1996.

[4] J. L. Bentley and A. C. Yao. An almost optimal algorithm for unbounded searching. *Info. Processing Letters*, 5:82–87, 1976.

[5] A. S. Besicovitch. On Kakeya's problem and a similar one. *Mathematische Zeitschrift*, 27:312–320, 1928.

[6] J. Canny and J. H. Reif. New lower bound techniques for robot motion planning problems. *IEEE Foundations of Computer Science*, 28:49–60, 1987.

[7] J. F. Canny, B. Donald, and E. Ressler. A rational rotation method for robust geometric algorithms. *Proc. 8th ACM Symp. on Computational Geometry*, pages 251–160, 1992. Berlin.

[8] J. Choi, J. Sellen, and C. Yap. Approximate Euclidean shortest path in 3-space. *Int'l. J. Computational Geometry and Applications*, 7(4):271–295, 1997. Also: Proc. 10th ACM Symp. on Comp. Geom., p.41–48, 1994.

[9] K. L. Clarkson. Approximationoalgorithms for shortest path motion planning, *Proc. 19th STOC*, 56–65, 1987.

[10] T. H. Corman, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms*. The MIT Press and McGraw-Hill Book Company, Cambridge, Massachusetts and New York, second edition, 2001.

[11] D. Halperin, L. Kavraki, and J.-C. Latombe. Robotics. In J. E. Goodman and J. O'Rourke, editors, *Handbook of Discrete and Computational Geometry*, chapter 41, pages 755–778. CRC Press LLC, 1997.

[12] J. Hershberger and L. J. Guibas. An $O(n^2)$ shortest path algorithm for non-rotating convex bodies. *J. Algorithms*, 9:18–46, 1988.

[13] C. Icking, G. Rote, E. Welzl, and C. Yap. Shortest paths for line segments. *Algorithmica*, 10:182–200, 1993.

[14] C. O'Dúnlaing, M. Sharir, and C. K. Yap. Retraction: a new approach to motion-planning. *ACM Symp. on Theory of Computing*, 15:207–220, 1983.

[15] J. O'Rourke. Finding a shortest ladder path: a special case. IMA Preprint Series 353, Institute for Mathematics and its Applications, University of Minnesota, 1987.

[16] C. H. Papadimitriou. An algorithm for shortest-path motion in three dimensions. *Inform. Process. Lett.*, 20:259–263, 1985.

[17] C. H. Papadimitriou and E. B. Silverberg. Optimal piecewise linear motion of an object among obstacles. *Algorithmica*, 2:523–539, 1987.

[18] J. T. Schwartz and M. Sharir. On the piano movers' problem: I. the case of a two-dimensional rigid polygonal body moving amidst polygonal barriers. *Communications on Pure and Applied Mathematics*, 36:345–398, 1983.

[19] J. T. Schwartz, M. Sharir, and J. Hopcroft, editors. *Planning, Geometry and Complexity of Robot Motion*. Ablex Series in Artificial Intelligence. Ablex Publishing Corp., Norwood, New Jersey, 1987.

[20] J. Sellen, J. Choi, and C. Yap. Precision-sensitive Euclidean shortest path in 3-Space. *SIAM J. Computing*, 29(5):1577–1595, 2000. Also: 11th ACM Symp. on Comp. Geom., (1995)350–359.

[21] M. Sharir. A note on the Papadimitriou-Silverberg algorithm for planning optimal piecewise-linear motion of a ladder. NYU Robotics Report 188, Courant Institute, New York University, 1989.

[22] M. Sharir, C. O'D'únlaing, and C. Yap. Generalized Voronoi diagrams for moving a ladder I: topological analysis. *Communications in Pure and Applied Math.*, XXXIX:423–483, 1986.

[23] M. Sharir, C. O'D'únlaing, and C. Yap. Generalized Voronoi diagrams for moving a ladder II: efficient computation of the diagram. *Algorithmica*, 2:27–59, 1987.

[24] Z. Sun and J. H. Reif. BUSHWHACK: An approximation algorithm for minimal paths through pseudo-Euclidean space. *Proc. 12th ISAAC*, Lecture Notes in Computer Science 2223, Springer-Verlag, 160–171, 2001.

[25] S. M. Ulam. *Problems of Modern Mathematics*. Science Editions, New York, 1964. Originally published as: *A Collection of Mathematical Problems*, Interscience Publishers, New York, 1960.

[26] C. K. Yap. Algorithmic motion planning. In J. T. Schwartz and C. K. Yap, editors, *Advances in Robotics, Vol. 1: Algorithmic and geometric issues*, chapter 3. Lawrence Erlbaum Associates, 1987.

[27] C. K. Yap. *Fundamental Problems in Algorithmic Algebra*. Oxford University Press, 2000. A version is available at URL `ftp:/Preliminary/cs.nyu.edu/pub/local/yap/algebra-bk`.

APPENDIX I: Basic Vocabulary

A **rod** is a closed line segment $\overline{AB}$, directed from the B-end (base) to the A-end (apex). The A-end is the end with the arrow head in figures. See Figure 1(a). So we also think of the rod as a line segment, directed from B-end to the A-end. The corresponding open segment is simply denoted as AB. The **focus** is a point F in the relative interior of the rod. The rod is thereby divided into two half-rods, AF and BF, viewed as open segments.

The closure and boundary of an arbitrary set $S \subseteq \mathbb{R}^2$ is denoted by $\overline{S}$ and ∂S, respectively. We are given a closed planar set $\Omega \subseteq \mathbb{R}^2$ in which the rod is free to move. Its boundary $\partial\Omega$ is polygonal, and is naturally partitioned into pairwise disjoint sets: singleton sets called **corners** and open line segments called **walls**. An **obstacle feature** refers to either a corner or a wall. The closure $\overline{W}$ of an open wall W is called a **closed wall**. For simplicity, we assume non-degeneracy on Ω as convenient.

We use the language of "placements" [26]. A **placement** is a pair $Z = (p, \theta) \in \mathbb{R}^2 \times \mathbb{S}^1$ where $p \in \mathbb{R}^2$ is a point and $\theta \in \mathbb{S}^1$ an angle. We also write $Z = (x, y, \theta)$ if $p = (x, y)$. For any set $S \subseteq \mathbb{R}^2$, we write $S[Z] \subseteq \mathbb{R}^2$ for **position** of the set S in placement Z. This position $S[Z]$ is obtained by

rotating the plane containing S about the origin by θ, then translating the plane by p, viewed upon some reference plane. For example, $\overline{AB}[Z]$ is the position of the rod in placement Z, and it is just a closed line segment. Thus the symbol "$[Z]$" acts as an Euclidean transformation of the plane. We choose a canonical representation of our rod AB: $F = (0,0)$, $A = (-a,0)$ and $B = (b,0)$ for some $0 < a < 1$ and $b = 1 - a$. For placement $Z = (x,y,\theta)$, we have $F[Z] = (x,y)$, $A[Z] = (x - a\cos\theta, y - a\sin\theta)$ and $B[Z] = (x + b\cos\theta, y + b\sin\theta)$. Furthermore,

$$\overline{AB}[Z] = \{(x + t\cos\theta, y + t\sin\theta : -a \leq t \leq b\}.$$

A placement Z is **free** if $\overline{AB}[Z] \subseteq \Omega$. Let $FP = FP(\Omega)$ denote the set of **free placements** (or configuration space). Consider a continuous function

$$\mu : [s,t] \to \mathbb{R}^2 \times \mathbb{S}^1$$

where $[s,t]$ is a real interval. For any point $X \in \mathbb{R}^2$, let $X\mu : [s,t] \to \mathbb{R}^2$ denote the path $X\mu(t) = X[\mu(t)]$. We call $X\mu$ the **X-trajectory** of μ. In case $X = F$ (the focus of the rod), the X-trajectory is called the **trace** of μ. We call μ a **potential motion** if both its A-trajectory and B-trajectory are rectifiable (i.e., has a definite arc length). This implies the trace of μ has a length $d_1(\mu)$, called the **d_1-distance** of μ. A potential motion μ is a **feasible motion** (or simply, "motion") if $\overline{AB}[\mu(u)] \subseteq \Omega$ for all $u \in [s,t]$. The d_1-distance between $Z, Z' \in FP$, denoted by $d_1(Z,Z')$, is the minimum d_1-distance of a feasible motion from Z to Z'. A motion μ is **d_1-optimum** if $d_1(\mu) = d_1(\mu(0), \mu(1))$.

For $r \geq 0$ and $p \in \mathbb{R}^2$, let $B_r(p)$ denote the closed Euclidean ball of radius r centered at p. Write $B_r(0)$ when p is the origin. For $S \subseteq \mathbb{R}^2$, let $B_r(S)$ denote the set $\cup\{B_r(p) : p \in S\}$ (alternatively, $B_r(S)$ is the Minkowski sum $B_r(0) \oplus S$). For a free placement Z, $B_r(AB[Z])$ has the shape of a capsule or "racetrack". The **clearance** Clearance(Z) is defined to be largest $r \geq 0$ such $B_r(AB[Z]) \subseteq \Omega$. Let the **racetrack** of Z refer to the set $B_r(AB[Z])$ when $r = h(Z)$; we denote the racetrack of Z by $RT(Z)$. We say Z is **closest** to those obstacle features s that intersect $RT(Z)$ on its boundary. The racetrack shown in Figure 3 shows two closest features, s_1, s_2. By definition, the interior of a racetrack has no obstacle points.

APPENDIX II: Structure of ∂FP

We focus on the boundary ∂FP of FP: ∂FP comprise those placements Z such that $\overline{AB}[Z]$ touches at least one obstacle feature. That is, $Z \in \partial FP$ iff there exists a wall W or a convex corner C such that at least one of three conditions hold: $A[Z] \in W$ or $B[Z] \in W$ or $C \in AB[Z]$. We will partition ∂FP into a cell complex.

Constraints. We describe the cell complex via the intermediate concept of **constraints** (the treatment here is slightly different from [3]). First, consider four types of **basic constraints**, denoted as

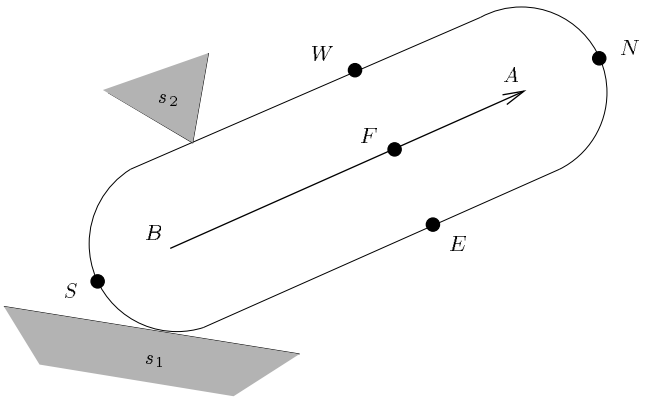


Figure 3: Racetrack $RT(Z)$: Closest features s_1, s_2 and Partition of its boundary

$$[B^+@W], [B^-@W], [B^+@C], [B^-@C].$$

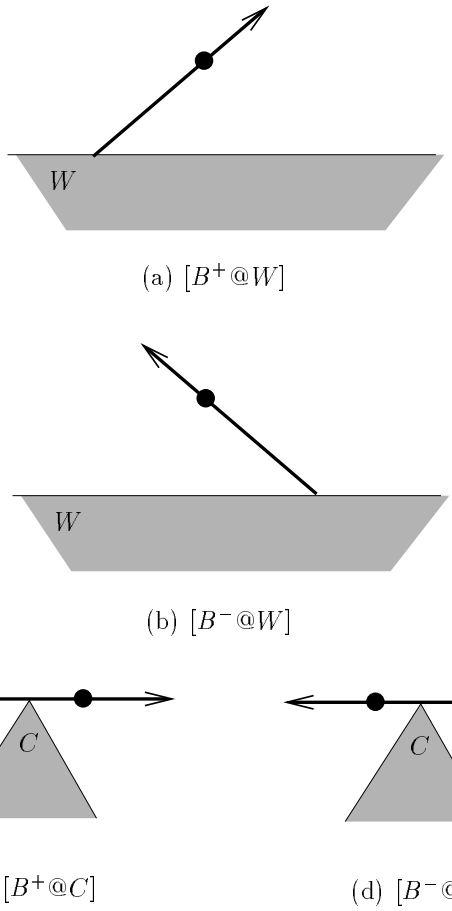
We could replace B by A to get 4 more basic constraints. Altogether there are 8 types of basic constraints. We emphasize that the corners C in constraints are always *convex* corners, i.e., the obstacle set $\mathbb{R}^2 \setminus \Omega$ is locally convex at C.

An obstacle feature s is a **CW stop** for Z if a CW rotation of the rod about the focus F, starting at Z, first becomes infeasible by virtue of crossing s. A **CCW stop** for Z is similarly defined. According to this definition, if an endpoint of a rod just grazes a feature s, then s is not considered a stop. For instance, in Figure 4(a), W is a CCW stop for the indicated placement Z.

Constraints are properties of placements. A placement Z with a constraint property ξ is said to **satisfy** that ξ, and written $Z \models \xi$. For instance, the constraint $[B^+@W]$ (read "B counter-clockwise at W") is satisfied by placement Z if $B[Z] \in W$ and W is a counter-clockwise (CCW) stop for Z (see Figure 4(a)). Similarly, $[B^-@C]$ ("B clockwise at C") is satisfied by Z if $C \in BF[Z]$ and C is a clockwise (CW) stop for Z (see Figure 4(d)). The other two basic constraints are illustrated in Figure 4(b) and (c).

Recall that BF is regarded as an open segment in this definition. Similarly, W is an open segment. A consequence of this definition is that if $B[Z] = C$, then Z does not satisfy any of the basic constraints. This is by design.

Patches, Edges, Vertices. For any basic constraint ξ, the set $\{Z \in FP : Z \models \xi\}$ is seen to be relatively open. Each connected component of this set is a 2-dimensional cell. The closure of such a cell is called a **ξ-patch**. Next, we define the 1-dimensional cells, also known as **edges** of FP. These are the connected components of the intersection of two patches. Again, we can use the intermediate concept of constraints: an **edge constraint** is any pair of basic constraints ξ_1, ξ_2, which we write as $\xi_1 \wedge \xi_2$. We say Z **satisfies** the edge



(a) $[B^+@W]$

(b) $[B^-@W]$

(c) $[B^+@C]$

(d) $[B^-@C]$

Figure 4: Basic Constraints

constraint, $Z \models \xi_1 \wedge \xi_2$, if[2] Z lies in the ξ_1-patch and the ξ_2-patch. For instance, $[B^-@C] \wedge [A^+@W]$ is an edge constraint satisfied by Z such that $C \in BF[Z]$, $A \in W[Z]$ and for which C and W are CW and CCW stops (respectively). If Z satisfies an edge constraint that has CW and CCW stops, we say Z is **constricted**. Finally the intersection of three or more *independent* patches is a 0-dimensional set. Each placement in such a set is called a **vertex**.

Note that by definition, patches and edges are closed sets: an edge contains vertices for its endpoints, and a patch contains edges for its boundary. Let an **open patch** be defined as a patch minus its boundary edges, and an **open edge** be an edge minus its endpoints. Then the open patches, open edges and vertices constitute a 2-complex for ∂FP.

[2]It important to realize that this is not the same as saying $Z \models \xi_1$ and $Z \models \xi_2$.

The classification of edges and vertices is more involved than for patches. Remark that in [3], we gave a somewhat different constraint analysis; in particular, the "edges" there do not coincide with the ones used here. But for our present purposes, it is enough to introduce three special classes of edges: mirrors, stopovers and pivots. Informally, the kind of edges excluded from this list are those satisfying an edge constraint which involve either two CW-stops or two CCW-stops.

Mirrors and Displaced Features. There are two kinds of mirrors: **straight mirrors** and **circular mirrors**. A placement Z of a mirror are said to be **reflecting**. The case of straight mirrors is easy to characterize: a straight mirror is an edge of ∂FP that satisfies a joint constraint $[X^+@W] \wedge [X^-@W]$, where $X = A$ or B and W is any wall. Figure 5(a) shows a wall W giving rise to two straight mirrors. But in general, due to the presence of other obstacles, a wall can give rise to a linear number of mirrors.

Corresponding to each mirror is a natural **mirror motion**; the trace of this motion is straight or circular, and is called a **displaced wall feature** (W_A, W_B in Figure 5(a)) or a **displaced corner feature** (C_A, C_B in Figure 5(b)). These displaced features are shown as dashed lines. Furthermore, the portion of the plane between the wall W and the displaced feature W_X ($X = A, B$) that is swept by the half rod FX is called the **zone** of the mirror.

Analysis of Circular Mirrors. We want to analyze the nature of the circular mirrors associated with a convex corner C. This analysis will also help clarify our decomposition of ∂FP based on patches. Consider the exterior angle at C, bounded by the incident walls W_1, W_2 (see Figure 5(c,d)). This angle is partitioned into 5 sectors, by extending the walls W_1, W_2 into the exterior angle, and by introducing normals to these walls at C. The extension and normal of W_i are denoted W_i^{ext} and $W_i^{\perp}$ in Figure 5(c) and (d).

Let M denote the sector bounded by the two normals; the rest of the exterior angle has two connected parts denoted E_1 and E_2. Let $FP_A(C) = \{Z \in FP : A[Z] = C, \overline{AB}[Z] \in M\}$ (there is a similar $FP_B(C)$). If $Z \in FP_A(C)$, we say Z **belongs** to a sector, say M, if $AB[Z]$ is contained in M, etc. We consider two cases, depending on whether the angle at C is greater than (Case (c)) or less than (Case (d)) a right angle.

- Case (c). In this case, E_i is in turn divided into two parts E_i' and E_i'' ($i = 1, 2$), where E_i' is the part adjacent to the wall W_i. Referring to Figure 5(c), note that a placement $Z \in FP_A(C)$ belongs to E_1' iff

$$Z \models [A^-@W_1] \wedge [A^-@C].$$

  Also, Z belongs to E_1'' iff

$$Z \models [A^-@W_1] \wedge [A^-@W_2].$$

  We replace W_1 by W_2 and A^- by A^+ in case of E_2' and E_2''. Finally, Z belongs to M iff

$$Z \models [A^-@W_1] \wedge [A^+@W_2].$$

- Case (d). In this case, M is divided into three parts named M_1, M_0, M_2. Referring to Figure 5(d), note that a placement $Z \in FP_A(C)$ belongs to E_1 iff

$$Z \models [A^- @ W_1] \wedge [A^- @ C].$$

In case of E_2, we replace A^- by A^+ and W_1 by W_2. Next, Z belongs to M_1 iff

$$Z \models [A^+ @ W_1] \wedge [A^- @ C].$$

Similarly, Z belongs to M_2 iff

$$Z \models [A^- @ W_2] \wedge [A^+ @ C].$$

Finally, Z belongs to M_0 iff

$$Z \models [A^+ @ W_1] \wedge [A^- @ W_2].$$

Summarizing cases (c) and (d), we can say $Z \in FP_A(C)$ belongs to M iff Z satisfies a joint clockwise (A^-) and counterclockwise (A^+) constraint.

Of course, the presence of other obstacles breaks up $FP_A(C)$ into several connected compoents. Each connected component in $FP_A(C)$, restricted to one of the 5 sector constitute an edge of ∂FP. By definition, a circular mirror is one of these edges comprising placements that belongs to M. The boundary between these sectors, when they are free, are vertices of ∂FP.

The terminology of "mirrors" comes from the fact that the trace of optimum motions sometimes "reflect" off the displaced wall or corner, the reflection obeying the law of reflection. This mirror phenomenon is critical for our NP-hardness proof in [3].

Stopover Edges and Stopover Curves. We define a **stopover edge** as a constricted edge whose trace is not straight. Recall that $Z \in \partial FP$ is **constricted** iff Z has a CW-stop and a CCW-stop, so it cannot rotate. The trace of a stopover edge is called a **stopover curve**. There are three kinds of stopover edges, and their constraints are:

(1) $[A^+ @ W_1] \wedge [B^- @ W_2]$ where W_1, W_2 are non-parallel walls: the trace is part of an ellipse.

(2) $[A^+ @ W] \wedge [X^- @ C]$ where $X = A$ or B: the trace is part of a conchoid.

(3) $[A^+ @ W_1] \wedge [A^- @ W_2]$: this can only happen when the A-end of the rod is (stuck) at a concave corner determined by W_1 and W_2. The trace is a circular arc.

Of course, in the above constraints, we could also interchange the roles of A and B. The last kind of constricted edge, which we do not consider to be a stopover edge, satisfies the constraints $[A^+ @ C_1] \wedge [B^- @ C_2]$ or $[A^+ @ C_1] \wedge [A^- @ C_2]$ or where C_1, C_2 are corners. The trace is a straight line segment.

Pivotal Edges. The third class of edges is very simple: they satisfy the **pivotal constraint** $[A^+ @ C] \wedge [B^- @ C]$ where C is any convex corner (the roles of A, B can be exchanged). Note that Z satisfies this constraint iff $F[Z] = C$.

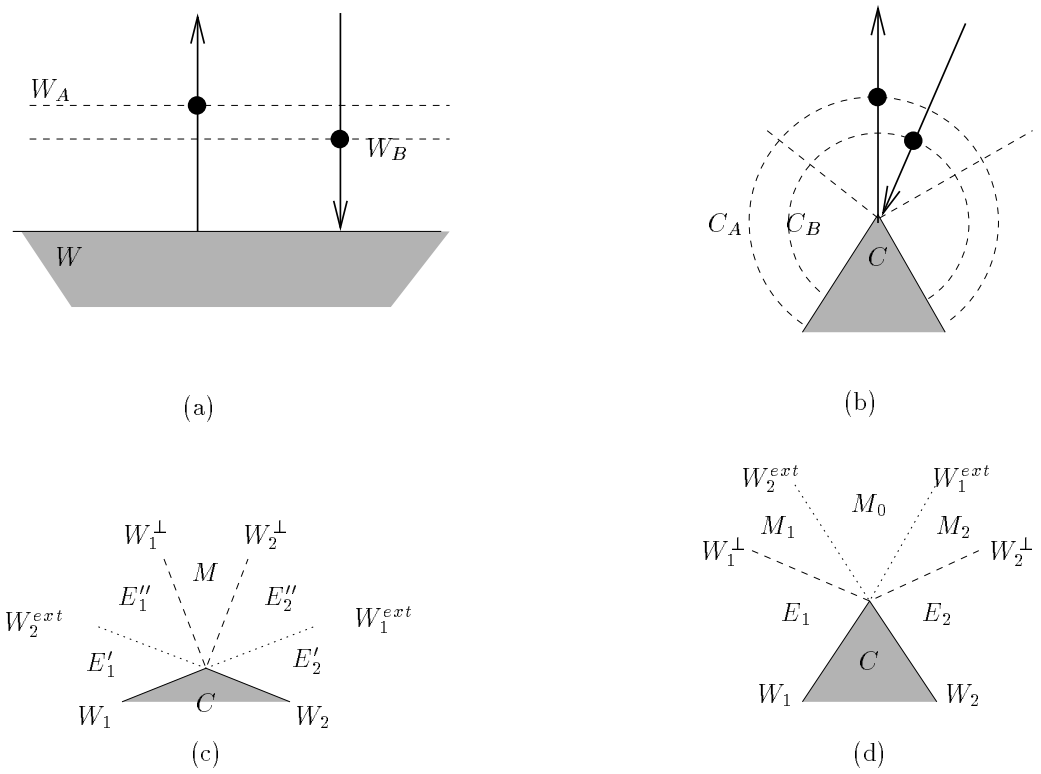


Figure 5: (a),(b) Mirrors and Displaced Features, (c),(d) Analysis of Circular Mirrors

We call Z **pivotal** in this case. A maximal connected set of such placements is a **pivotal edge** and corresponds to a motion of pure rotation about the focus which is fixed at some C.

Complexity of the Cell Complex. We conclude this appendix by bounding the size of the cell complex of ∂FP which we just described.

THEOREM 8 *Let there be n wall and corner features. In our 2-complex of ∂FP, the following bounds hold:*
(a) The number of patches is $O(n^2)$.
(b) The number of edges is $O(n^3)$.
(c) The number of vertices is $O(n^3)$.

Proof. (a) Each patch is determined by one of the $8n$ basic constraints. A basic constraint ξ can give rise to $< n$ patches. To see this suppose the constraint ξ relates to a wall W, say $\xi = [A^+@W]$. In the absense of any other features

there is only one patch. We now introduce features one at a time: each feature can only increase the number of patches by 1. First we introduce the corner features. Each corner feature, if it increases the number of patches, can only do so by split one current patch into two new patches. Now we introduce the wall features (their corners have already been introduced). But these cannot increase the number of patches. A similar argument applies when ξ relates to a corner C, as in $\xi = [A^+@C]$.

(b) There are $64n^2$ edge constraints. Each edge constraint $\xi_1 \wedge \xi_2$, in the absence of any other features, determines a 1-dimensional set of placements that is a connected set. Again, each new feature we introduce, if it disrupts the connectivity of any connected component, can only increase the number of components by 1. Hence there are at most $64n^3$ edges in the complex.

(c) Each vertex can be charged to a triple of obstacle features. There are n^3 such triples. But every triple can be charged a constant number of times. Hence number of charges is $O(n^3)$. **Q.E.D.**

APPENDIX III: Proof of Local Characterization Theorem

We prove Theorem 5 which gives a local characterization the d_1-optimal motion.

The following classification of placements is based on considering the behavior of the clearance function as we make small rotatations, where rotations are always about the focus F:

DEFINITION 1 *(1) The placement Z is* **stopped** *if* Clearance(Z) *is locally maximum in the sense that any infinitesimal rotation (CW- or CCW-) will decrease the clearance.*
(2) A placement Z is **critical** *if* Clearance(Z) *is locally minimum in the sense that any infinitesimal rotation (CW- or CCW-) will increase the the clearance.*
(3) Define a **rotation function** *$\rho : FP \to FP$ as follows. If Z is stopped or critical, let $\rho(Z) = Z$. Otherwise, there is a unique direction to rotate Z so as to locally increase the clearance of Z. Define $\rho(Z)$ to be the first local maxima reached by this rotation. We also define*

$$H(Z) = \text{Clearance}(\rho(Z)).$$

In Appendix I, we introduced the racetrack $RT(Z)$. See Figure 3. If Clearance(Z) > 0, then the boundary of $RT(Z)$ contains four special points called the **North**, **South**, **East**, **West Poles** of Z. The boundary is thereby divided into four open curves, called the **North-East**, **North-West**, **South-East**, **South-West tracks**. Stopped placements can be characterized as placements whose East or West Poles are covered, or two adjacent tracks are covered. (Two tracks are adjacent if they are both bounded by one of the four poles. E.g., the North-East track is adjacent to North-West and to South-East tracks.) Similarly, we may characterize critical placements to be those placements Z whose

North or South Poles are "covered" (i.e., contained in some obstacle feature), and furthermore, *no point on the boundary of $RT(Z)$ is covered.* The latter condition is important, as it implies that a critical placements could not be simultaneously stopped.

These characterizations depend upon the assumption Clearance$(Z) > 0$, but they can be extended to the case Clearance$(Z) = 0$ by taking limits. Thus, stopped placements and critical placements are generalizations (respectively) of constricted placements and reflecting placements. The requirement that critical placements must not have any other points on its racetrack boundary covered is translated into the requirement that a reflecting placement must not satisfy any other constraints (other than what is required by definition).

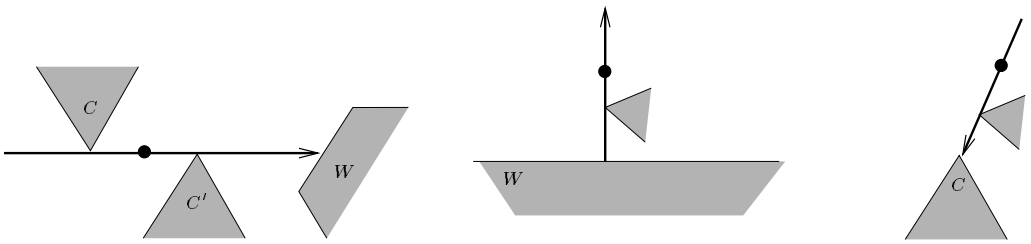
Next we introduce some special motions. Let $\mu : [0, 1] \to FP$ be a motion.

- μ is **constricted** if for all $t \in [0, 1]$, $\mu(t)$ is constricted.
- μ is **pivotal** if for all $t \in [0, 1]$, $\mu(t)$ is pivotal. In this case, there is a unique convex corner C such that $F\mu(t) = C$.
- μ is **reflecting** if for all $t \in [0, 1]$, $\mu(t)$ is reflecting.
- μ is **restrained** if it is constricted, pivotal or reflecting.
- A placement Z is **restrained** if it is one of the following: (a) constricted, (b) pivotal, (c) reflecting, or (d) a vertex. REMARK: The clause (d) might seem unnecessary, but Figure 6(i) shows a vertex V_1 that does not fall under (a), (b) or (c). More important is the remark that our definition of a reflecting placement Z does not allow the Z to satisfy any other constraints. Thus the vertices V_2 and V_3 in Figure 6(ii,iii) are also not covered by clauses (a), (b) or (c). In proofs, we do not care if the vertex V_1 is considered restrained (though it is harmless to be considered restrained); but we do need V_2, V_3 to be considered restrained.
- A motion μ is **unrestrained** if for all t, $0 < t < 1$, $\mu(t)$ is not restrained. Note $\mu(0)$ and $\mu(1)$ may be restrained in this definition.
- μ is **straight** if for all $t \in [0, 1]$, $F\mu(t)$ lies on the straight line segment $[F\mu(0), F\mu(1)]$. Moreover, the trace is monotone in the sense that for $0 \leq t < t' \leq 1$, $F\mu(t')$ is closer than $F\mu(t)$ to $F\mu(1)$.

. We now prove two lemmas about unrestrained placements.

LEMMA 9 *If Z is unrestrained, then $H(Z) > 0$.*

Proof. If Clearance$(Z) > 0$, then the result follows from the fact that $H(Z) \geq$ Clearance(Z). So assume Clearance$(Z) = 0$. In this case, $\overline{AB}[Z]$ must touch some feature s (wall or corner). There is a unique direction in which we rotate Z away from s; this is because s cannot be a corner at $F[Z]$ (it would make Z pivotal), and s is not a wall normal to $AB[Z]$ (it would make Z reflecting). Moreover, if we make a small enough rotation, Z would remains free (otherwise Z is constricted). This proves that $\rho(Z) \neq Z$, *i.e.*, $H(Z) > 0$. **Q.E.D.**



(i) V_1 (ii) V_2 (iii) V_3

Figure 6: Vertices that are not constricted, pivotal or reflecting

LEMMA 10 *Let $\mu : [0,1] \to FP$ be an unrestrained motion. If μ is d_1-optimal, then μ is straight.*

Proof. Suppose for all $0 < t < 1$, Clearance$(\mu(t)) > 0$. Then it is an easy remark that μ must be straight.

Our goal is to construct a new motion μ' with three properties: (1) Clearance$(\mu'(t)) > 0$, (2) $\mu'(t) = \mu(t)$ for $t = 0, 1$, and (3) μ, μ' have the same trace, $F\mu' = F\mu$.

¿From (3), we have $d_1(\mu') = d_1(\mu)$ (= length of their traces). Since μ is optimal, and from (2), we conclude that μ' must be optimal. From (1) and the easy remark, μ' must be straight. Using (3) again, it follows that μ is also straight, proving our lemma.

Hence it remains to construct μ'. Towards this end, consider $\mu'' = \rho \circ \mu$ (composition of ρ with μ). We see that $F\mu'' = F\mu$. Furthermore, Lemma 9 tells us that $H(\mu(t)) > 0$ for all $0 < t < 1$. Thus Clearance$(\mu''(t)) = H(\mu(t)) > 0$. So, μ'' has properties (1) and (3) needed for μ'. We can get (2) by concatenating a rotation at the beginning and at the end of μ''. So what else is lacking? Unfortunately, μ'' can have discontinuities: this happens at those $0 < t_0 < 1$ where $\mu(t_0)$ is critical. By Sard's Lemma [18, 19], we may assume that there only finitely many such discontinuities, which must be isolated. Now μ' is obtained from μ'' by "patching up" each of these discontinuities via a rotation at each discontinuity.

To see how this patchwork is achieved, observe that an isolated discontinuity at t_0 means that $\mu''(t_0^-) \neq \mu''(t_0^+)$. Moreover, $\mu''(t_0^+)$ and $\mu''(t_0^-)$ must be the two local maximas for the clearance function as we rotate $\mu(t_0)$ in the CW- and CCW- directions. We can connect $\mu''(t_0^-)$ to $\mu''(t_0^+)$ by a rotation motion that passes through $\mu(t_0)$. The minimum clearance achieved by this rotation motion is equal to Clearance$(\rho(\mu(t_0))) =$ Clearance$(\mu(t_0)) > 0$. In other words, by inserting such rotation motions into μ'', we preserve the property that Clearance$(\mu'(t)) > 0$. This completes the description of μ'. **Q.E.D.**

We next show two more preliminary results about constricted motions. If

$p \in \mathbb{R}^2$ and $\varepsilon > 0$, let $B_\varepsilon(p) \subseteq \mathbb{R}^2$ denote the **open ball** of radius $\varepsilon > 0$ centered at $p \in \mathbb{R}^2$. For $Z \in FP$, define the **reachable ball** $B_\varepsilon(Z) \subseteq FP$ comprising those placements Z' that can be reached by a motion $\mu : [0,1] \to FP$ where the trace $F\mu$ is restricted to $B_\varepsilon(F[Z])$. Define the set of **ε-accessible points** from Z to be $\{F[Z'] : Z' \in B_\varepsilon(Z)\}$, and denoted $FB_\varepsilon(Z)$. Clearly, $FB_\varepsilon(Z) \subseteq B_\varepsilon(F[Z])$. Also, $B_\varepsilon(F[Z]) \setminus FB_\varepsilon(Z)$ is the set of **ε-inaccessible points** from Z.

LEMMA 11 *Let $\mu : [0,1] \to FP$ be a constricted edge. Then the trace $F\mu$ is the boundary of the locally accessible points. That is, for any $0 < t < 1$, there is an $\varepsilon > 0$ such that*
(i) The ε-accessible points $FB_\varepsilon(\mu(t))$ and the ε-inaccessible points $B_\varepsilon(F\mu(t)) \setminus FB_\varepsilon(\mu(t))$ are connected sets.
(ii) The restriction of $F\mu$ to $B_\varepsilon(\mu(t))$ is a connected curve that separates $FB_\varepsilon(\mu(t))$ from $B_\varepsilon(F\mu(t)) \setminus FB_\varepsilon(\mu(t))$.

Proof. As illustrated in Figure 7, the trace $F\mu$ is one of four types: (a) is an eliptic arc, (b) is a circular arc, (c) and (d) represent upper and lower conchoid arcs, while (e) and (f) are straight segments. To see that the trace $F\mu$ forms the boundary for the locally accessible points, we note that for any $0 < t < 1$, there is an open range $R = R(t)$ of angles at $F\mu(t)$ such that for any $\theta \in R$, there is a feasible motion starting from $\mu(t)$ with trace moving in the direction θ. Moreover, every motion starting from $\mu(t)$ in the opposite direction $-\theta$ is infeasible. **Q.E.D.**

Next we analyze how a restrained motion and a constricted motion can be joined together in an optimal motion. We first introduce the necessary notation. Let $\mu : [a,b] \to FP$ and $\mu' : [a',b'] \to FP$ be two motions. We can **concatenate** them provided $b = a'$ and $\mu(b) = \mu'(a')$ and obtain a new motion denoted $\mu'' : [a,b'] \to FP$ where $\mu''(t) = \mu(t)$ if $t \in [a,b]$ and $\mu''(t) = \mu'(t)$ otherwise. We also write $\mu'' = \mu;\mu'$ for the concatenation of μ and μ'.

LEMMA 12 *Let $\mu = \mu_0;\mu_1$ be d_1-optimal with μ_0 constricted and μ_1 unrestrained. Let Z be the placement where μ_0 joins μ_1. Assume $F\mu_0$ and $F\mu_1$ are both non-constant functions and the set of ε-inaccessible points from Z is convex. Then $F\mu_0$ is straight and connects to $F\mu_1$ tangentially at the point $F[Z]$.*

Proof. Since μ_1 is unrestrained, its trace $F\mu_1$ is straight. By Lemma 11, $F\mu_0$ locally bounds the inaccessible points. Thus the tangent lines of μ_1 (locally) lies in the accessible region. Let $F\mu_0$ meet $F\mu_1$ at the point r. See Figure 8(a). If $F\mu_1$ is not a tangent to $F\mu_0$ at r, we can construct a shorter motion to obtain a contradiction. To do this, take a neighborhood $B_\varepsilon(r)$ as in Lemma 11. Join a point p on the $F\mu_0$ to a point q on $F\mu_1$ by a line segment within this neighborhood. This segment is in the locally accessible region and μ can be modified to take a shortcut by tracing this segment. **Q.E.D.**

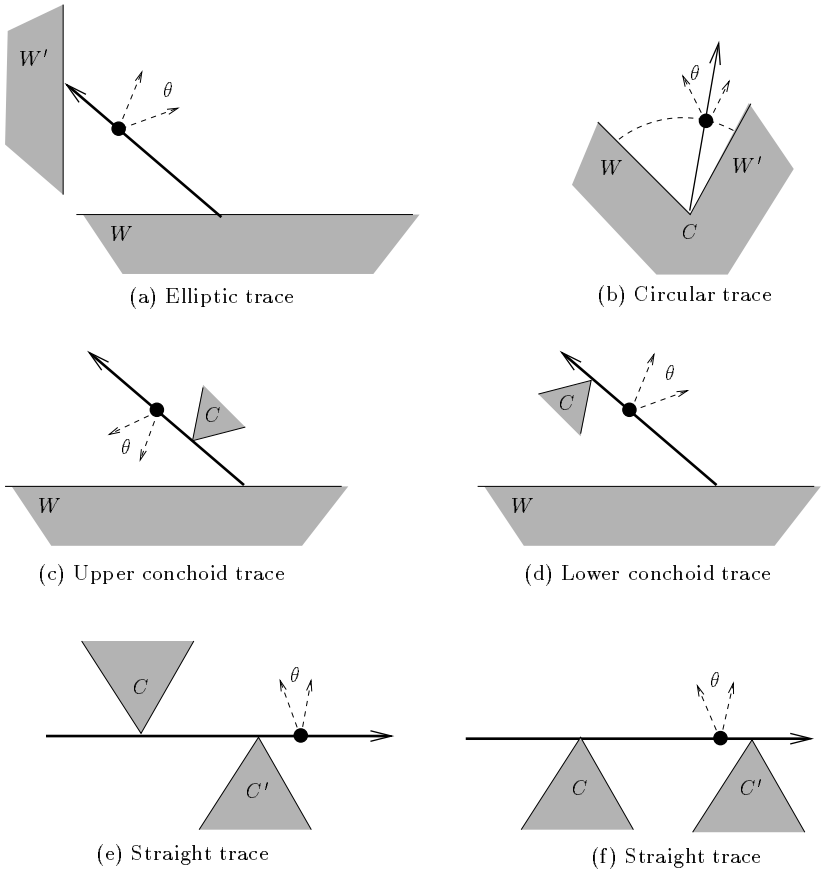


Figure 7: Local Geometry of Constricted Motions

The set of ε-inaccessible points from a non-vertex constricted Z is generally convex for ε sufficiently small (being bounded by straight segments or arcs of ellipses, circles or conchoids). Unfortunately, there is an exception: when the arc is the part of a conchoid that is closer to the directrix of the conchoid than to the pole. In this case the ε-accessible points need not be convex. This is illustrated in Figure 8(b). We treat this case next.

LEMMA 13 *Let $\mu = \mu_0; \mu_1$ be a motion with μ_0 constricted and μ_1 unrestrained. Let Z be the placement where μ_0 joins μ_1. Assume $F\mu_0$ and $F\mu_1$ are both non-constant functions and for all $\varepsilon > 0$, the set of ε-accessible points from Z is non-convex. Then μ is not optimal.*

Proof. If μ is optimal, then μ_1 is straight. Moreover, the trace of μ_0 is part of a conchoid curve γ (see Figure 8(b)). We can choose a point $p = F\mu_0(t_0)$ and a point $q = F\mu_1(t_1)$ (for suitable t_0, t_1), both sufficiently close to $F[Z]$,

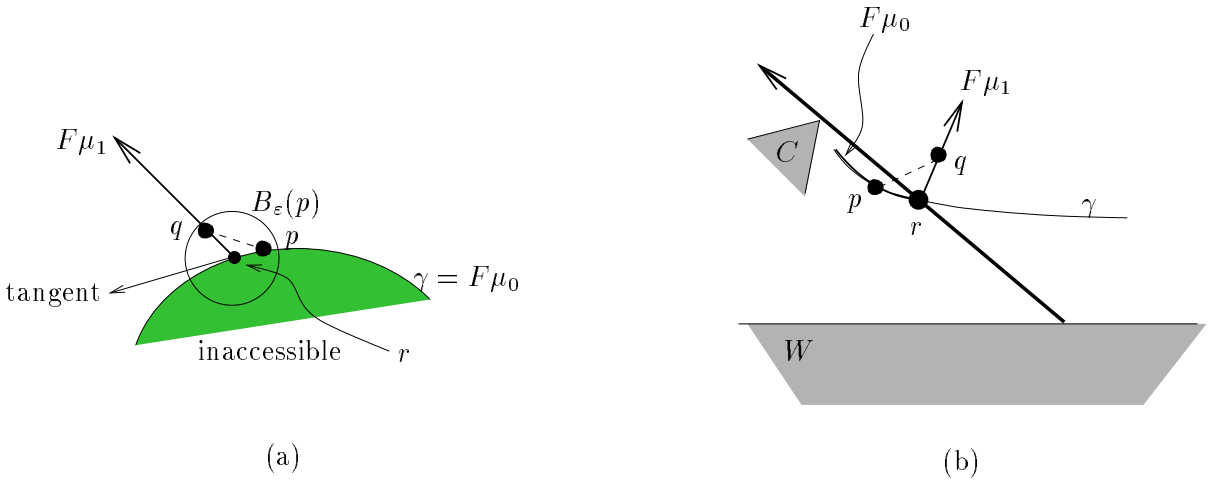


Figure 8: Transition between constricted motion μ_0 and unrestrained motion μ_1: (a) convex case, (b) nonconvex case

such that there exists a straight motion from $\mu_0(t_0)$ to $\mu_1(t_1)$. This proves the non-optimality of μ. **Q.E.D.**

Proof of the Local Characterization (Theorem 5): Let $F\mu$ be locally non-straight at t_0, $0 < t_0 < 1$. If $\mu(t_0)$ is locally a vertex at t_0, then we satisfy the first condition in the theorem. In the rest of this proof, we assume otherwise.

CLAIM: μ is locally restrained at t_0, i.e., if I is any $F\mu$-neighborhood of t_0, there exists some $t \in I$ such that $\mu(t)$ is restrained.

By way of contradiction, assume the claim is false. Then Lemma 10 tells us that μ restricted to I is straight. This contradicts our assumption that μ is not locally straight at t_0.

¿From this claim, we conclude that μ must be locally constricted, pivotal or reflecting at t_0. We consider each possibility in turn.

(i) Suppose μ is locally pivotal at t_0. That it is clear that $\mu(t_0)$ is in fact pivotal, i.e., there is a convex corner C such that $F\mu(t_0) = C$. Let $I = [t_1, t_2]$ be the maximal interval containing t_0 such that $F\mu(t) = C$ for all $t \in I$. Since t_0 is not locally a vertex, it means that there exists a $\varepsilon > 0$ such that $\mu|[t_1 - \varepsilon, t_1]$ and $\mu|[t_2, t_2 + \varepsilon]$ is unrestrained and thus straight. Moreover, the trace of these two straight motions must "bend" around C because of optimality.

(ii) Suppose μ is locally constricted at t_0. Again, we conclude that $\mu(t_0)$ is constricted. So let $I = [t_1, t_2]$ be the maximal interval containing t_0 such that $F\mu(t)$ is constricted for all $t \in I$. We have two possibilities: (a) If $t_0 = t_1$

or $t_0 = t_2$, then $\mu(t_0)$ is the transition between a constricted motion $\mu|[t_1, t_2]$ and some other motion, say μ'. Since $\mu(t_0)$ is not locally a vertex, we conclude that μ' cannot be restrained. By Lemmas 12 and 13, we conclude that $F\mu'$ is straight, and meets the trace $F\mu_i$ tangentially at $\mu(t_0)$. (b) If $t_1 < t_0 < t_2$, then $\mu(t_0)$ is locally tracing a stopover curve.
(iii) Finally, assume $\mu(t_0)$ is locally reflecting. Let $I = [t_1, t_2]$ be the essential $F\mu$-neighborhood of t_0. Then there is some $t_1 \in I$ where $\mu(t_3) \in M$, where M is a mirror. Without loss of generality, let $A[\mu(t_3)]$ lie in some wall or corner. According to our analysis of mirrors in Appendix II, $\mu(t_3)$ satisfies some edge constraint of the form $[A^+@s] \wedge [A^-@s']$ where s, s' are two features. There are three possibilities (a) $s = s'$ are the same wall, (b) s is a corner and s' is an incident wall (or vice-versa), (c) s, s' are the two walls incident on a common corner. For all sufficiently small $\varepsilon > 0$, the points $F\mu(t_1 - \varepsilon)$, $F\mu(t_3)$, $F\mu(t_2 + \varepsilon)$ are not collinear. This is just our assumption that μ is not locally straight at t_0. Let $R \subseteq \mathbb{R}^2$ denote the zone of the mirror M – this is the region $\cup\{AF[Z] : Z \in M\}$. Now, if $F\mu(t_1 - \varepsilon)$ or $F\mu(t_2 + \varepsilon)$ lies outside the zone R, then it is not hard to see that we can define a straightline motion from $\mu(t_1)$ to $\mu(t_2)$. This is a shortcut, contradicting the d_1-optimality of μ. So both points lies in the zone. This means that $\mu(t_1 - \varepsilon)$ satisfies $[A^+@s]$ and $\mu(t_2 + \varepsilon)$ satisfies $[A^-@s']$ (the symmetric case is treated similarly). In this case, the optimal motion from $\mu(t_1 - \varepsilon)$ to $\mu(t_2 + \varepsilon)$ must have a trace that reflects off the mirror curve of M according to Snell's law, as claimed.

This completes the proof of Theorem 5.