Discrete Comput Geom 2:255–270 (1987)

Discrete & Computational Geometry

On the Number of Critical Free Contacts of a Convex Polygonal Object Moving in Two-Dimensional Polygonal Space*

Daniel Leven[1] and Micha Sharir[1,2]*

[1] School of Mathematical Sciences, Tel-Aviv University, Tel-Aviv, Israel

[2] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA

Abstract. We show that the number of critical positions of a convex polygonal object B moving amidst polygonal barriers in two-dimensional space, at which it makes three simultaneous contacts with the obstacles but does not penetrate into any obstacle is $O(kn\lambda_s(kn))$ for some $s \le 6$, where k is the number of boundary segments of B, n is the number of wall segments, and $\lambda_s(q)$ is an almost linear function of q yielding the maximal number of "breakpoints" along the lower envelope (i.e., pointwise minimum) of a set of q continuous functions each pair of which intersect in at most s points (here a breakpoint is a point at which two of the functions simultaneously attain the minimum). We also present an example where the number of such critical contacts is $\Omega(k^2n^2)$, showing that in the worst case our upper bound is almost optimal.

1. Introduction

Let B be a convex polygonal object having k vertices and edges, and free to move (translate and rotate) in a closed two-dimensional space V bounded by a collection of polygonal obstacles ("walls") having altogether n corners. The problem studied in this paper is to estimate the number of *free critical* positions of B at which it makes simultaneously three distinct contacts with the walls. More precisely, a free position is one at which B is fully contained in V, so that it can touch some obstacles, but not penetrate into any of them (this notation differs from that of [7], where such positions are called *semifree*). A critical free

* Work on this paper by the second author has been supported by Office of Naval Research Grant N00014-82-K-0381, National Science Foundation Grant No. NSF-DCR-83-20085, and by grants from the Digital Equipment Corporation, and the IBM Corporation.

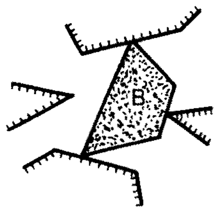

Fig. 1.

position Z is one at which there exist three distinct pairs (W_1, S_1), (W_2, S_2), (W_3, S_3), such that for each $i = 1, 2, 3$ either W_i is a wall segment and S_i is a corner of B or W_i is a wall corner and S_i is a side of B, and such that at position Z S_i touches W_i. See Fig. 1 for an illustration of such a critical contact. Note that this somewhat "liberal" definition also regards as critical, positions at which a corner of B touches a wall corner and another corner/side of B touches another wall edge/corner (including also placements at which a side of B overlaps a wall segment with an endpoint in common).

This problem is a major subproblem in the design and analysis of efficient algorithms for automatic planning of a continuous obstacle-avoiding motion of B within V between two specified positions. If B is a line segment (a "ladder"), then it is shown in [5] that the total number of such critical positions of B is $O(n^2)$, which consequently leads to an $O(n^2 \log n)$ algorithm for the desired motion planning. If B is a convex polygonal object which is free only to translate in V but not to rotate, then the motion-planning problem becomes simpler and can be accomplished in time $O(n \log n)$ [4], [6]. This follows from the property, proved in [4] and related to the problem studied in the present paper, that the number of free positions of B (all having the same given orientation) at which it simultaneously touches *two* obstacles is only $O(n)$ (provided the obstacles are in "general position" [4]). If B is also allowed to rotate then, extending the motion-planning technique of [5], one obtains an algorithm whose complexity depends on the number of critical free positions of B at which it makes simultaneously three distinct contacts with the walls. Since each such contact is a contact of either a corner of B with a wall edge or of an edge of B with a wall corner, a crude and straightforward upper bound on the number of these critical positions of B is $O((kn)^3)$. Moreover, if B is nonconvex, then there are cases where the number of these critical positions of B is indeed $\Omega((kn)^3)$ (cf. a remark in [4]). However, in this paper we show that if B is convex, then the number of these critical triple contacts is only $O(kn\lambda_s(kn))$, for some $s \le 6$, where $\lambda_s(q)$ is the maximal number of "breakpoints" along the lower envelope (i.e., pointwise minimum) of a set of q continuous functions, each pair of which intersect in at most s points. Here a breakpoint is a point at which two of the functions simultaneously attain the minimum; note that between any two adjacent breakpoints the minimum is attained by a single function, thus $\lambda_s(q)+1$ is equal to the maximal number of connected graph portions composing the lower envelope of a collection of such functions. It is shown in [10] that $\lambda_s(q) = O(q \log^* q)$

(where $\log^* q$ is the length of the smallest exponential tower $2^{2^{\cdot^{\cdot^{2}}}}$ exceeding q, and where the constant of proportionality depends on s). A better asymptotic bound is given in [2] for the case $s=3$ and in [8] for larger values of s. These better bounds are roughly of the form $\lambda_s(q)=O(q\alpha(q)^{O(\alpha(q)^{s-3})})$, where $\alpha(q)$ is the functional inverse of Ackermann's function, and is thus extremely slowly growing. In short, for a fixed s, $\lambda_s(q)$ is nearly linear in q, although, as has been shown in [2] and [9], $\lambda_s(q)=\Omega(q\alpha^{\lfloor(s-1)/2\rfloor}(q))$, so that it is superlinear in q for $s\geqq 3$.

Finally, using these results, and modifying the technique of [5] to the case of an aribtrary convex polygonal moving body B, one can obtain an $O(kn\lambda_s(kn)\log(kn))$ motion-planning algorithm for B. This algorithm will be described in full detail in a forthcoming companion paper [3].

2. Estimating the Number of Critical Free Contacts

In this section we prove our main result, namely (see Theorem 2.4) that the number of critical free contacts of B, as defined in the introduction, is $O(kn\lambda_s(kn))$ for some $s\leq 6$. We assume that B and the set of obstacles are in *general position.* By this we mean that the shape of B and the positions of the obstacles are such that there does not exist a position of B at which it satisfies four independent constraints imposed on B by its possible contacts with obstacles. Each constraint of this form either requires a specific corner of B to touch a specific wall edge, or requires a specific side of B to touch a specific wall corner, or requires that the segment connecting two points of contact be perpendicular to the wall edge or to the side of B involved in one of these contacts. Furthermore, these constraints must be *independent,* in the sense that none of them is a consequence of the others in an arbitrary generic placement of B and the obstacles; for example, for a contact of a corner S of B against a wall corner W, only two out of the four possible constraints that can be associated with this contact (namely, S touching the two wall edges incident to W and W touching the two sides of B incident to S) can be independent, as is easily checked. Typical configurations in which the obstacles and B are not in general position are shown in Fig. 2; they involve respectively, from left to right, (i) four distinct points of simultaneous contact, (ii) three points of contact where one is of a corner of B against a wall corner (thus contributing two independent constraints), and (iii) two points of contact

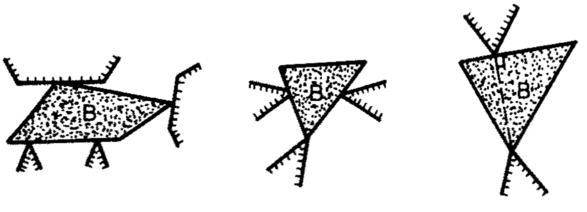

Fig. 2.

where one is a corner–corner contact and where the segment connecting these points is perpendicular to a side of B.

Definition 2.1. (a) A (potential) *contact pair* O is a pair (W, S) such that either W is a (closed) wall edge and S is a corner of B or W is a wall corner and S is a (closed) side of B. The contact pair is said to be of type I in the first case, and of type II in the second case.

(b) An actual *obstacle contact* (i.e., a contact of B with an obstacle) is said to *involve* the contact pair $O=(W, S)$ if this contact is of a point on S against a point on W, and, furthermore, if this contact is *locally free*, i.e., the inner angle of B at S lies entirely on the exterior side of W if S is a corner of B, and the entire angle within the wall region V^c at W lies exterior to B if W is a wall corner.

(c) The *tangent line* T of a contact pair $O=(W, S)$ is either the line passing through W if W is a wall edge or the line passing through W and parallel to S if S is a side of B (in the second case T depends of course on the orientation of B).

It is clear from the above definition that there are $O(kn)$ possible contact pairs. Throughout this section we will use the same index i to refer to a contact pair O_i, to its corresponding wall element W_i and boundary element S_i of B, and to its tangent line T_i. Let Z be a free position of B at which it makes two simultaneous obstacle contacts involving the contact pairs O_i, O_j for which T_i, T_j are not parallel. Then $z_{ij}=z_{ji}$ will denote the intersection point of T_i and T_j, x_{ij} (resp. x_{ji}) will denote the contact point of W_i with S_i (resp. W_j with S_j) at position Z, and l_{ij} will denote the line passing through x_{ij} and x_{ji}. Also u_{ij}, v_{ij} will denote the endpoints of W_i if O_i is of type I or the endpoints of S_i when B is positioned at Z if O_i is of type II, such that u_{ij} and z_{ij} lie on the same side of x_{ij} on T_i, and v_{ij} lies on the other side of x_{ij} (u_{ji}, v_{ji} are defined similarly for O_j) (see Fig. 3).

An extreme situation arises when $x_{ij}=z_{ij}$, which is the case if W_i is a wall edge, W_j is a wall corner, and S_i is an endpoint of the side S_j of B. In this case u_{ij} and v_{ij} are not well defined, although u_{ji} and v_{ji} are. We will refer to this case by calling O_i, O_j *adjacent* contact pairs. Another extreme situation arises when a corner S of B touches a wall corner W; formally speaking (see also the discussion at the beginning of this section), this can be regarded as a double contact (e.g., of S against the two wall edges meeting at W). In this case we have $x_{ij}=x_{ji}=z_{ij}$, so that all four points u_{ij}, v_{ij}, u_{ji}, v_{ji}, as well as the line l_{ij}, are not

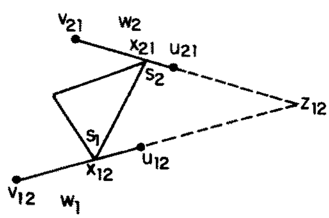

Fig. 3.

well defined. Although we will use such double contacts in our lower-bound example given at the end of the paper, we will ignore such singular contacts in our upper-bound analysis (as justified in a remark preceding Theorem 2.4).

Definition 2.2. Let O_1, O_2 be two contact pairs. We say that O_2 *bounds* O_1 at the orientation Θ if there exists a (not necessarily free) position $Z = (X, \Theta)$ of B at which it makes two simultaneous obstacle contacts involving O_1, O_2, respectively, such that $B^* = \mathrm{conv}(S_1 \cup S_2)$ always intersects W_2 as we move B from Z without changing the orientation Θ along the tangent T_1 in the direction of z_{12} until the last position at which S_1 still touches W_1. (In case of adjacent contact pairs O_1, O_2 with O_1 of type I and O_2 of type II, we have $z_{12} = x_{12}$ so that the direction of motion of B in this definition is not defined; moreover, in this case B^* is simply the segment S_2, which stops intersecting W_2 immediately as we move it toward either endpoint of W_1. In this ill-defined case we prefer to regard O_2 as not bounding O_1 (note, however, that in this case O_1 does bound O_2 in accordance with the above definition).)

The crucial property on which our analysis depends is:

Proposition 2.1. *Let O_1, O_2 be two contact pairs for which there exists a position $Z = (X, \Theta)$ of B at which it makes two simultaneous obstacle contacts involving O_1, O_2, respectively. Then either O_1 bounds O_2 at Θ or O_2 bounds O_1 at Θ, except when the corresponding tangents T_1, T_2 are coincident or parallel. (Typical degenerate cases of kind are:* (a) *W_1, W_2 are both wall corners and $S_1 = S_2$;* (b) *W_1, W_2 are both wall corners and S_1, S_2 are parallel sides of B;* (c) *W_1, W_2 are parallel wall edges; and* (d) *W_1* is a wall edge and *S_2 is a side of B* parallel to *W_1*.)

Proof. In the nondegenerate case the tangents T_1, T_2 must intersect at a single point. Using the terminology introduced above, we consider separately three possible subcases (see Fig. 4).

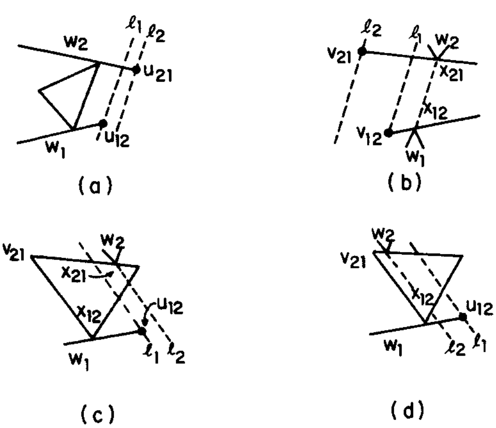

Fig. 4.

(1) W_1, W_2 are both wall edges (Fig. 4(a)).
Consider the line l_1 (resp. l_2) parallel to l_{12} passing through u_{12} (resp. u_{21}). Obviously either l_1 intersects W_2 or l_2 intersects W_1. Suppose, without loss of generality, that l_1 intersects W_2. But then as we move B along W_1 from x_{12} until S_1S_2 reaches l_1, the segment $B^* = S_1S_2$ will always intersect W_2. Thus O_2 bounds O_1 at Θ.

(2) W_1, W_2 are both wall corners (Fig. 4(b)).
Let l_1 (resp. l_2) be the line parallel to l_{12} which passes through v_{12} (resp. v_{21}). Obviously either l_1 intersects S_2 or l_2 intersects S_1. Suppose, without loss of generality, that l_1 intersects S_2. Then, since B^* is convex, the entire trapezoid Δ bounded by l_{12}, l_1, S_1, and S_2 is contained in B^* (in this position of simultaneous contact with W_1 and W_2). Thus as we move B along T_1 toward z_{12} until the endpoint v_{12} of S_1 coincides with W_1, then throughout this motion W_2 will intersect Δ, hence B^*, again showing that O_2 bounds O_1.

(3) W_1 is a wall edge and W_2 is a wall corner (or symmetrically, W_2 is a wall edge and W_1 is a wall corner) (Fig. 4(c) and (d)).
If O_1 and O_2 are adjacent contact pairs, then obviously O_1 bounds O_2 (see a comment in Definition 2.2). Assume that this is not the case and let l_1 (resp. l_2) be the line parallel to S_1v_{21} passing through u_{12} (resp. through $x_{21} = W_2$). Obviously, either l_1 intersects the line segment $x_{21}v_{21}$ or l_2 intersects the line segment $x_{12}u_{12}$. But in the first case, observe that at the position of mutual contact the triangle $x_{12}x_{21}v_{21}$ is contained in B^*; thus if we move B along W_1 until S_1 meets u_{12}, then throughout this motion W_2 must intersect this triangle, hence B^*, so that in this case O_2 bounds O_1. In the second case if we translate B along T_2 until v_{21} meets W_2, then throughout this motion the line segment S_1v_{21} (and hence also B^*) will intersect W_1 so that in this case O_1 bounds O_2. □

Remark. The proof of Proposition 2.1 actually implies that, for each double contact of B involving pairs O_1, O_2, there exists a critical orientation $\Theta_{O_1O_2}$ of B (for which the lines l_1 and l_2 coincide) such that at orientations greater than $\Theta_{O_1O_2}$ one of the contact pairs, say O_1, bounds O_2 but O_2 does not bound O_1, and at orientations smaller than $\Theta_{O_1O_2}$ the contact pair O_2 bounds O_1 but O_1 does not bound O_2. (If O_1, O_2 are adjacent, then on one side of $\Theta_{O_1O_2}$ one of these contact pairs, say O_1, always bounds O_2, whereas on the other side of $\Theta_{O_1O_1}$ no such double contact of B is possible.)

Let O_1 be any contact pair and consider all contact pairs that bound O_1 (at any orientation Θ). For each such pair O_2 we define the function $F_{O_1O_2}(\Theta)$ over the domain $\Pi = \Pi_{O_1O_2}$ of orientations Θ of B in which O_2 bounds O_1, to be the distance of x_{12} from v_{12} at the position $Z = (X, \Theta)$ in which B simultaneously makes two obstacle contacts involving O_1, O_2, respectively (see Fig. 5; note that v_{12} is always well defined whenever O_2 bounds O_1).

Note that Π need not in general be connected but may consist of several subintervals, such that, for orientations Θ outside these intervals, either two obstacle contacts involving O_1, O_2, respectively, cannot occur simultaneously, or O_2 does not bound O_1. Nevertheless we have:

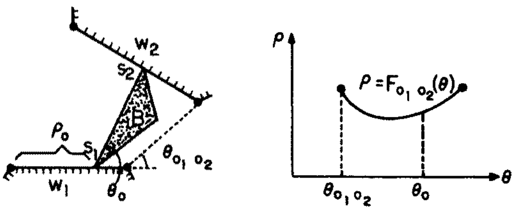

Fig. 5.

Lemma 2.2. $\Pi_{O_1O_2}$ *consists of at most five intervals.*

Proof. O_2 bounds O_1 at an orientation Θ of B if and only if the following four conditions hold:

(1) Θ lies on the appropriate side of the critical orientation $\Theta_{O_1O_2}$ which separates between the domain in which O_2 bounds O_1 and the domain in which O_1 bounds O_2 (see the remark following Proposition 2.1).
(2) There exists at orientation Θ a position of B at which S_1 touches W_1 and S_2 touches W_2 simultaneously.
(3) Suppose S_1 is a corner of B and W_1 is a wall. Then at this position of double contact, the inner angle of B at S_1 lies entirely on the exterior side of W_1; if S_1 is a side of B and W_1 is a wall corner, then again the entire angle within the wall region V^c at W_1 should lie exterior to B. Similar conditions should hold for the contact of S_2 against W_2.
(4) The tangents T_1, T_2 of O_1, O_2, respectively, are not parallel or coincident at Θ.

Thus the domain of definition of the function $F_{O_1O_2}$ is the intersection of the three domains I_1, I_2, I_3 satisfying, respectively, conditions (1), (2), and (3) above. (Condition (4) pre-empts, regardless of the value of Θ, cases in which W_1 and W_2 are parallel or coincident, or S_1 and S_2 are parallel or coincident, or W_1 and W_2 are two wall corners touching the same side of B, etc. However, if the contact pairs O_1, O_2 are of "mixed" types, e.g., O_1 is of type I and O_2 is of type II, then condition (4) defines an orientation Θ^* at which S_2 becomes parallel to W_1; since at this orientation the corner v_{12} involved in the definition of $F_{O_1O_2}$ changes discontinuously, we artificially split the domain of $F_{O_1O_2}$ at Θ^* into two disjoint subdomains, and seek to prove the assertion of the lemma for each of these subdomains separately.)

Note first that the domain I_3 is always an angular interval. Note also that, since the angular range for Θ is circular, the domain I_1 is not well defined. However, combining condition (1) with condition (3), we can define I_1 as an angular interval, one of whose endpoints is $\Theta_{O_1O_2}$, and the other is the appropriate endpoint of I_3. In conclusion, $I_1 \cap I_3$ is simply an angular interval.

The domain I_2 of Θ satisfying (2) has a more complicated structure. The following arguments provide a crude analysis of this structure.

Suppose first that W_1, W_2 are both wall edges. Then $B^* = \mathrm{conv}(S_1 \cup S_2)$ is a straight segment. By continuity, if at Θ the segment B^* touches W_1, W_2 at interior points of these walls, and if, as we assume, W_1 and W_2 are not parallel, then B^* will continue to have a simultaneous contact with W_1, W_2 at all orientations at a sufficiently small neighborhood of Θ. Hence, to find orientations on the boundary of I_2. it suffices to consider orientations at which either S_1 touches an endpoint of W_1 or S_2 touches an endpoint of W_2 at positions of B in which it makes the two contacts involving the pairs O_1, O_2 simultaneously. Each one of these four kinds of touch can clearly occur in at most two orientations of B (corresponding to the two points at which the circle of radius $|S_1S_2|$ about the common corner of contact intersects the other wall edge). Thus altogether I_2 has at most eight boundary orientations, and since it lies in the circular range of Θ, it follows that I_2 consists of at most four disjoint angular intervals. (Although this is not essential for subsequent analysis, it should be observed that our assumption on general position of the obstacles rules out cases in which some of the eight critical orientations just defined are isolated points of I_2, or are interior points of I_2.)

Next suppose that W_1, W_2 are both wall corners. In this case we can use the fact that the role of the boundary sides and corners of B and the role of the wall edges and corners are completely symmetrical in condition (2), so that, applying symmetric arguments to those used above, one can show that in this case too I_2 consists of at most four intervals. (A similar idea is also used in the proof of Lemma 2.3 below.)

Finally, consider the case in which W_1 is a wall edge and W_2 is a wall corner (or the other way around). Here it can also be seen that if B makes at orientation Θ a simultaneous double contact involving the pairs O_1, O_2 so that S_1 touches an interior point of W_1 and W_2 touches an interior point of S_2, then such a simultaneous double contact would also be possible at all orientations sufficiently close to Θ. Thus similar arguments to those used above imply that in this case too I_2 consists of at most four disjoint intervals.

Hence the required domain $\Pi = I_1 \cap I_2 \cap I_3$ consists of at most five intervals, as can be easily checked. This completes the proof of the lemma. □

If $\Pi_{O_1O_2}$ is indeed not connected we will consider each connected portion of $F_{O_1O_2}$ as a separate partially defined function. Clearly, the number of such functions is still at most $O(kn)$, for each fixed contact pair O_1.

Next recall that the definition of v_{12} (and thus also of $F_{O_1O_2}$) depends on the position of the intersection point of T_1, T_2 relative to the contact point of S_1 with W_1, and that v_{12} is one of the endpoints of W_1 if O_1 is of type I or one of the endpoints of S_1 if O_1 is of type II. Consequently, we partition the collection of "bounding functions" $F_{O_1O_2}$ for O_1 into two classes A_1, A_2 so that for all functions in A_1, v_{12} is the same endpoint of W_1 (or of S_1), whereas for all functions in A_2 it is the other endpoint. (Thus if O_1, O_2 are contact pairs of mixed types, the domain of $F_{O_1O_2}$ has to be split at the orientation Θ^* at which the side of B involved in one contact becomes parallel to the wall edge involved in the other contact, so that on one side of Θ^* the function $F_{O_1O_2}$ is in A_1, whereas on the other side of Θ^* it belongs to A_2.) Thus each contact pair O_1 defines two "complementary" coordinate frames (Θ, ρ) which can be used to represent positions of B at which it makes an obstacle contact involving O_1. Here Θ is the orientation of B and ρ is the distance between the contact point of the contact involving O_1 and a designated endpoint v of either W_1 or S_1. Within each frame, let C_{O_1} denote the polar domain representing positions of B at which a contact involving O_1 occurs. Consider one such coordinate frame for O_1 and the corresponding collection A_1 of bounding functions (for which the endpoint v_{12} coincides with the designated endpoint v for that coordinate frame). It follows from definition that if $F_{O_1O_2} \in A_1$, $(\Theta, \rho) \in C_{O_1}$, $\rho_0 = F_{O_1O_2}(\Theta)$ is defined, and $\rho > \rho_0$, then (Θ, ρ) is a nonfree position of B (because W_2 intersects the interior of B in this position); in other words, the area of C_{O_1} above each bounding function in A_1 represents nonfree positions of B.

Our goal is to estimate the number of *critical free* positions of B at which it makes three simultaneous obstacle contacts, or, more precisely, as defined in the introduction. Let $Z = (X, \Theta)$ be such a position and let the corresponding contact pairs involved in the simultaneous contacts be $O_i = (W_i, S_i)$, $i = 1, 2, 3$. For each pair $i \neq j = 1, 2, 3$, either O_i bounds O_j or O_j bounds O_i at the orientation Θ of B (unless one of the degenerate situations listed in Proposition 2.1 occurs). Hence, in general, the position Z is represented by points lying on the graphs of some of the bounding functions $F_{O_iO_j}$. Suppose, without loss of generality, that it lies on $F_{O_1O_2} \in A_1$. Then, since Z is free, we must have

$$F_{O_1O_2}(\Theta) = \min\{F_{O_1O}(\Theta)\colon F_{O_1O} \in A_1\};$$

in other words, Z is represented by a point on the *lower envelope* $\phi_{O_1A_1}$ of the functions in A_1. (Note that the converse does not necessarily hold, i.e., a position represented by a point lying on $\phi_{O_1A_1}$ may be nonfree, because B might intersect, at this position, obstacles whose contacts with B involve pairs that do not bound O_1 at Θ.) Furthermore, since at Z the object B also makes an obstacle contact involving the pair O_3, we must have one of the following situations:

(i) O_3 also bounds O_1 at Θ and $F_{O_1O_3}$ (over some neighborhood of Θ) also belongs to A_1. In this case Z is represented by an intersection point of $F_{O_1O_2}$ and $F_{O_1O_3}$ on $\phi_{O_1A_1}$ (we will refer to such a point as a *break-point* of $\phi_{O_1A_1}$) (Fig. 6(a)).

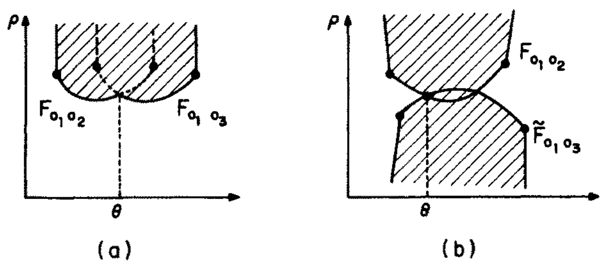

Fig. 6. Schematic illustration of the representation of various types of critical contact.

(ii) O_3 also bounds O_1 at Θ, but $F_{O_1O_3}$ (over some neighborhood of Θ) belongs to A_2. Let $\tilde{\phi}_{O_1A_2}$ denote the lower envelope of the functions in A_2, reflected into the coordinate frame (Θ, ρ) in which the functions in A_1 are represented (since this transformation is a reflection of the ρ-axis, $\tilde{\phi}_{O_1A2}$ will be the upper envelope of the reflections $\tilde{F}_{O_1O}$ of the functions F_{O_1O} in A_2). Then in this case Z is represented as an intersection point of $\phi_{O_1A_1}$ with $\tilde{\phi}_{O_1A_2}$ (Fig. 6(b)).

(iii) O_3 does not bound O_1 at Θ, and moreover, no two of these contact pairs simultaneously bound the third pair at Θ, but two of these pairs, say O_1, O_3, have parallel or coincident associated tangents (Fig. 7(a)).

(iv) As in (iii), no two of these contact pairs simultaneously bound the third one, and further, the degenerate cases in (iii) above do not arise. In this case we can assume, without loss of generality, that O_1 does not bound O_2, O_2 does not bound O_3, and O_3 does not bound O_1 at Θ. But then by Proposition 2.1 it must be the case that O_2 bounds O_1, O_3 bounds O_2, and O_1 bounds O_3 at Θ, so that Z is represented by a point on $F_{O_1O_2}$, a point on $F_{O_2O_3}$ and a point on $F_{O_3O_1}$, each lying in a corresponding lower envelope (Fig. 7(b)).

We now proceed to estimate separately the number of critical contacts of each of these four types.

Type (i) Critical Contacts

Let O_1 be a contact pair and consider the lower envelope $\phi = \phi_{O_1A_1}$ of one of the collections A_1 of bounding functions for O_1. To estimate the number of breakpoints along ϕ, we first extend each function in A_1 over the complete range of Θ as follows (see also [1]). Suppose a function $F_{O_1O_2} \in A_1$ is defined over an interval (θ_1, θ_2) of Θ. We then extend $F_{O_1O_2}$ leftward from θ_1, in a continuous manner along a ray of slope $-K$ and similarly extend it rightward from θ_2 along a ray of slope $+K$ for some large $K > 0$. Obviously, if we choose K uniformly for all functions in A_1 to be sufficiently large, then the sequence of functions of A_1 as they appear along the lower envelope ϕ will not change under this extension. Therefore if one can show that each pair of (extended) functions in A_1 intersect in at most some fixed number s of points, then by the results reviewed in the

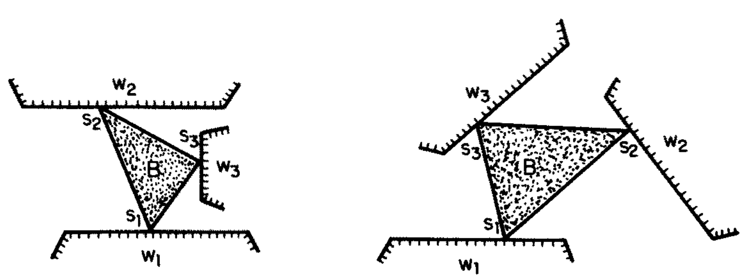

Fig. 7. Critical contacts of types (iii) and (iv).

introduction, the number of such intersections lying on the graph of ϕ is bounded by $\lambda_s(O(kn)) = O(\lambda_s(kn))$.

To show that each pair of our functions does indeed intersect in at most some fixed number s of points, we use the following lemma, and the observations that each intersection of two (unextended) functions $F_{O_1O_2}$, $F_{O_1O_3} \in A_1$ represents a position of B in which it simultaneously makes three contacts involving the pairs O_1, O_2, O_3, respectively, and that the extended portions of these functions can contribute at most two additional intersection points for each such pair of functions.

Lemma 2.3. *Let O_1, O_2, O_3 be three distinct contact pairs. Then there are at most four positions of B at which it makes simultaneously three contacts involving O_1, O_2, O_3, respectively.*

Proof. By the analysis in [7] it follows that the curve $\gamma_{O_1O_2}$ traced by some reference point on B as B makes simultaneously two obstacle contacts involving the pairs O_1, O_2 is either a straight segment, or part of an ellipse, or part of a quartic algebraic curve. Hence, if one of the contact pairs, say O_3, is of type I then we can take S_3 to be the reference point on B, so that the desired positions of triple contact of B correspond to the intersection points of $\gamma_{O_1O_2}$ with the straight segment W_3. Since $\gamma_{O_1O_2}$ is at most quartic, there are at most four such intersections, thus at most four positions of B at which this triple contact can occur. If all three contact pairs are of type II then we can consider a coordinate frame in which B is stationary and the obstacles move in a collectively rigid manner. Then the desired positions of triple contact correspond to positions in this coordinate frame of the moving triangle $W_1W_2W_3$ at which its vertices touch simultaneously the three segments S_1, S_2, S_3, respectively. Hence this case can be treated in much the same way as the case in which all contact pairs are of type I, and the preceding argument implies that there are at most four positions of triple contact in this case too. □

Remark. If all contact pairs are of type I (or, symmetrically, all are of type II) then (see [7]) $\gamma_{O_1O_2}$ is either a straight segment or part of an ellipse, so that in these cases there exist at most two positions of such a triple contact. In the two remaining cases of triple contact, the curves $\gamma_{O_iO_j}$ are more complex and we do not know whether the number of desired positions of triple contact is also two in these cases. We conjecture that this is indeed the case.

Hence the number of points of intersection of each pair of extended functions $F_{O_1O_2}$, $F_{O_1O_3}$ is at most six (or four if the above conjecture is true). Thus, summing the number of such critical break-points, over all contact pairs O_1, we conclude that there are at most $O(kn\lambda_s(kn))$ critical contacts of type (i), for some fixed (but unfortunately still unknown) constant $s \leq 6$.

Remark. Note that Lemma 2.3 states that there are at most four positions of B at which it makes a triple contact involving the contact pairs O_1, O_2, O_3. However, it is not clear whether all these positions can appear on the same lower envelope for O_1, so that it is conceivable that the actual s might even be smaller than the above estimate.

Type (ii) Critical Contacts

Let O_1 be a contact pair. We wish to estimate the number of intersections of the lower envelope $\phi_{O_1A_1}$ with the reflected envelope $\tilde{\phi}_{O_1A_2}$. For this, consider the collection of all orientations Θ at which either $\phi_{O_1A_1}$ or $\tilde{\phi}_{O_1A_2}$ has a break-point. Clearly, there are $O(\lambda_s(kn))$ such orientations which partition the angular range of Θ into $O(\lambda_s(kn))$ disjoint intervals. Let I be one of these intervals. Then there are two contact pairs O_2, O_3 such that $\phi_{O_1A_1} \equiv F_{O_1O_2}$ and $\phi_{O_1A_2} \equiv F_{O_1O_3}$ on I. Hence each intersection point of the envelopes $\phi_{O_1A_1}$ and $\tilde{\phi}_{O_1A_2}$ within I is an intersection of $F_{O_1O_2}$ and the reflection of $F_{O_1O_3}$. But since each such intersection corresponds to a position of B at which it makes simultaneously three contacts involving O_1, O_2, O_3, Lemma 2.3 and the discussion following it imply that the total number of intersections of these functions is at most $s \leq 6$. Thus on each such interval I, $\phi_{O_1A_1}$ and $\tilde{\phi}_{O_1A_2}$ intersect only in $O(1)$ points, so that the total number of their intersection points is at most $O(\lambda_s(kn))$. Summing over all contact pairs O_1, it follows that there are at most $O(kn\lambda_s(kn))$ critical contacts of type (ii).

Type (iii) Critical Contacts

Let Θ be an orientation of B at which it makes simultaneously three contacts involving the pairs O_1, O_2, O_3, such that no two of them bound simultaneously the third pair at Θ, and such that two of these contact pairs, say O_1, O_2, satisfy at Θ one of the degenerate conditions listed in Proposition 2.1. Since we assume that B and the obstacles are in general position, it can be checked that none of these degenerate cases can arise (at Θ) also for O_1, O_3 or for O_2, O_3. Since by assumption O_1, O_2 do not both bound O_3, it then follows that O_3 bounds one of them, say O_1. Hence this critical contact is represented as a point on one of the envelopes for O_1. But it is easily seen that there are only $O(kn)$ orientations Θ of B in which it can make a contact involving O_1 and another contact involving O_2 in one of those degenerate manners. Moreover, since the critical type (iii) contact that we consider must be represented by a point lying on an envelope for O_1, it follows that each such Θ can determine only one critical contact of the above form, represented by the point at orientation Θ on that envelope of O_1 that coincides with $F_{O_1O_3}$ at Θ. Hence, summing over all possible contact pairs O_1, we obtain at most $O(k^2n^2)$ (which is also $O(kn\lambda_s(kn))$) critical contacts of type (iii).

Type (iv) Critical Contacts

Finally, consider the case of type (iv) contacts. Consider first the set C of all critical orientations at which some envelope ϕ_{OA} has a break-point. Without loss of generality we can assume that each $\Theta \in C$ is defined by a unique triple of contact pairs. (Otherwise, if some $\Theta \in C$ is induced by more than one triple of contact pairs, then, applying an infinitely small perturbation to the obstacle configuration, we can split Θ into several orientations, infinitely close to one another, each now induced by a unique triple of contact pairs; see [7] for details on this perturbation technique.) By the preceding arguments, C consists of $O(kn\lambda_s(kn))$ orientations, which partition the angular range for Θ into $O(kn\lambda_s(kn))$ disjoint noncritical intervals. Consider one such interval I. For each contact pair O_1, each of the envelopes $\phi_{O_1A_1}$, $\phi_{O_1A_2}$ is equal over I to a single bounding function in A_1, A_2, respectively. Suppose I contains an orientation Θ_0 at which a type (iv) critical contact occurs, which involves O_1 and two additional contact pairs O_2, O_3. Also, without loss of generality, assume that O_2 bounds O_1, O_3 bounds O_2, and O_1 bounds O_3 at Θ_0. Then throughout I, one of the lower envelopes for O_1 coincides with $F_{O_1O_2}$, one of the lower envelopes for O_2 coincides with $F_{O_2O_3}$, and one of the lower envelopes for O_3 coincides with $F_{O_3O_1}$. It follows that to find all possible triples O_1, O_2, O_3 of contact pairs which include a specific contact O_1, and which can induce a type (iv) free contact at some orientation within I, one simply has to consider the two contact pairs whose bounding functions appear on the lower envelopes for O_1 over I, then obtain, for each of these contact pairs O_2, the two contact pairs representing the two lower envelopes for O_2 over I, and finally check that O_1 is a contact pair representing one of the envelopes for such a third contact pair O_3. Hence there exist at most four such triples of contact pairs (involving a specific O_1), so that altogether the lower envelopes over I induce at most $O(kn)$ critical orientations at which a type (iv) contact can occur. (Note that not all these induced orientations necessarily lie in I; but even if such an orientation Θ lies outside I, it can still realize the corresponding free type (iv) critical contact, because the functions appearing in the corresponding lower envelopes over I may still appear there also over the noncritical interval containing Θ.)

Now let I' be an interval adjacent to I, and let Θ^* be their common endpoint. Θ^* is a critical orientation at which, by assumption, B makes a unique triple obstacle contact involving three contact pairs O_1^*, O_2^*, O_3^*. This implies that as we cross through Θ^* from I to I', only the functions appearing in the lower envelopes for O_1^*, O_2^*, O_3^* can change. But the preceding arguments then implies that only $O(1)$ new critical contacts of type (iv) can be induced by the various lower envelopes over I', in addition to those that were already induced by the envelopes over I. In other words, each noncritical interval can contribute only $O(1)$ additional potential contacts of type (iv), so that altogether there can be at most $O(kn\lambda_s(kn))$ critical contacts of type (iv).

Remark. So far we have been ignoring critical contacts involving a contact pair O_1 of a corner S_1 of B against a wall corner W_1 (since O_1 involves two independent constraints on the position of B, we seek here critical contacts involving O_1 and just one more contact pair O_2). Such double critical contacts however are quite easy to analyze. Indeed, for such a contact pair O_1, the only degree of freedom left for B as it makes the contact involving O_1 is rotation about the common point of contact S_1 against W_1. During this rotation, each additional contact involving some pair $O_2 = (W_2, S_2)$ can occur in at most two orientations, so that there are only $O(kn)$ potential critical orientations at which a critical contact involving O_1 can occur. Hence, sorting these orientations in circular order, and processing them one at a time in this order, it is easy to determine which of these orientations yields a free critical contact of the sort we seek; we leave details of this straightforward procedure to the reader. Thus altogether there are at most $O(k^2n^2)$ critical contacts of this form, and they can all be calculated in time $O(k^2n^2 \log(kn))$.

Similar remarks apply to the case in which a side S_1 of B overlaps a wall edge W_1. Again this condition leaves only one degree of freedom to vary (namely that of sliding S_1 along W_1), and one can show in the same manner as above that only $O(kn)$ critical contacts involving O_1 are possible, and that they can all be found in $O(kn \log(kn))$ time. Thus there are at most $O(k^2n^2)$ singular contacts of this second kind, and they can all be found in time $O(k^2n^2 \log(kn))$.

All this gives us the following main theorem.

Theorem 2.4. *The number of critical free triple contacts of a convex k-sided polygonal object B moving amidst polygonal obstacles composed of n walls is* $O(kn\lambda_s(kn))$ *for some* $s \leq 6$.

Remark. The preceding analysis can be used to obtain an efficient procedure for calculating all critical contacts of B. Roughly, this procedure first calculates all lower envelopes for contact pairs. This can be done, using a divide-and-conquer approach, in time $O(kn\lambda_s(kn)\log(kn))$ as outlined, e.g., in [2]. Then type (ii) and type (iv) critical contacts can be calculated by a "sweeping process" which iterates over the noncritical intervals of Θ, and maintains a priority queue of potential critical contacts, which requires a constant number of updates as we cross from one noncritical interval to an adjacent one. Hence, altogether this process also requires $O(kn\lambda_s(kn)\log(kn))$ time. Type (iii) critical contacts are even easier to calculate. Hence all critical free contacts of B can be calculated in time $O(kn\lambda_s(kn)\log(kn))$. More details on this procedure, and on the overall efficient motion-planning algorithm which is based on this procedure, is given in a forthcoming companion paper [3].

3. A Lower Bound on the Number of Critical Contacts

We conclude this paper by giving an example in which the number of critical free contacts of a convex k-gon B with a collection of polygonal obstacles having n corners altogether, is $\Omega(k^2n^2)$, thus showing that, in the worst case, our upper bound is close to optimal.

Let n and k be given. The moving body B is defined as follows. Let z be the center of a circle with radius l_1, and let $x_1, \ldots, x_k$ be k equally spaced points arranged in counterclockwise order along an arc of that circle. The angle $\phi = \angle x_i z x_{i+1}$, for all $i = 1, \ldots, k-1$, is chosen so that $k\phi \ll \pi/n$. Since $d(z, x_i x_{i+1}) = l_1 \cos(\phi/2) < l_1$, for each $i = 1, \ldots, k-1$, it follows that for sufficiently small $\varepsilon_2 > \varepsilon_1 > 0$ there exists a sufficiently small $\delta_0 > 0$ so that for all points y at distance $l_2 < \delta_0$ from z we have

$$d(y, x_j) > l_1 - \varepsilon_1 > l_1 - \varepsilon_2 > d(y, x_j x_{j+1}) \qquad (*)$$

for all j.

Let $y_1, \ldots, y_k$ be k points arranged in counterclockwise order along a circle of radius l_2 about z, for some $l_2 < \delta_0$, so that they all lie on the convex hull $\operatorname{conv}\{x_1, \ldots, x_k, y_1, \ldots, y_k\}$. Define B to be that hull, so that B is a convex $(2k)$-gon (see Fig. 8(a)).

Next we define the obstacles to consist of just the following $2n$ corners (see Fig. 8(b)): The first n corners $u_1, \ldots, u_n$ are taken to be points equally spaced (at angles π/n apart) on a semicircle centered at the origin O whose radius is $r_1 = l_1 - (\varepsilon_2 + \varepsilon_2)/2$. There exists $\delta > 0$ such that for each point w at distance δ from O we have

$$l_1 - \varepsilon_1 > d(w, u_i) > l_1 - \varepsilon_2$$

for all $i = 1, \ldots, n$. The last n corners $v_1, \ldots, v_n$ are then defined so that each v_i lies on the segment Ou_i at distance δ from O.

We claim that if we choose $\delta_0 \le \pi\delta/n$, then for each obstacle corner v_j, $j = 1, \ldots, n$, and for each pair of a segment $x_p x_{p+1}$, $p = 1, \ldots, k-1$, and a corner y_q, $q = 1, \ldots, k_1$, there are $\Omega(n)$ obstacle corners u_i for which there exists a free position of B at which it makes the two contacts of $x_p x_{p+1}$ with u_i and of y_q with v_j simultaneously.

Indeed, place y_q at v_j and rotate B about this common point of contact. It is easy to see that, by the choice of ϕ and δ_0, B can be rotated in this way almost 180° without touching any other corner v_i. Thus, during this rotation B will meet

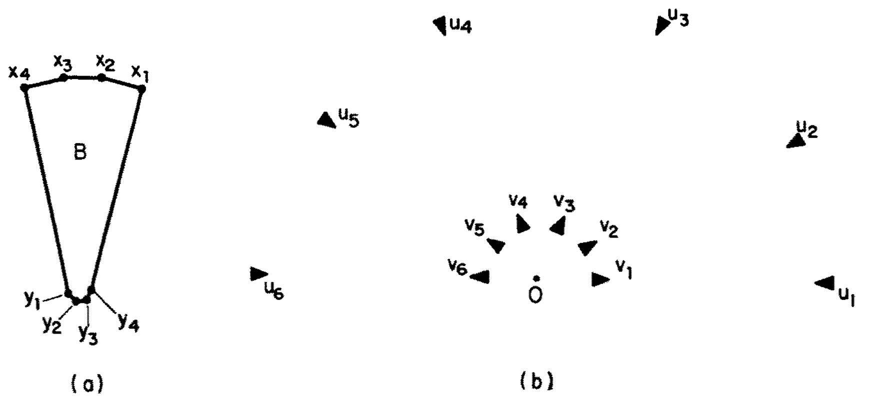

Fig. 8.

$\Omega(n)$ corners u_i, and our choice of ϕ implies that, at any position during this rotation, B cannot meet two such corners simultaneously. Finally, it follows from the inequalities (*) that every side $x_p x_{p+1}$ of B will touch such a corner u_i as B rotates in this manner, and that every such contact will necessarily be at a free position of B.

Hence in this example we have $\Omega(k^2 n^2)$ distinct free double contacts of B with the obstacles, each involving three simultaneous *contact pairs*, namely the contacts of $x_p x_{p+1}$ with u_i, of $y_{q-1} y_q$ with v_j, and of $y_q y_{q+1}$ with v_j, thus yielding the desired lower bound.

Remark. We do not have a similar example in which a convex k-gon B makes $\Omega(k^2 n^2)$ free critical contacts with a collection of polygonal obstacles having n corners altogether, in such a way that in each of these critical contacts B touches the obstacles at three distinct points. It is easy, however, to obtain (by modifying the above example or otherwise) examples in which B makes $\Omega(kn^2)$ (or, symmetrically, $\Omega(k^2 n)$) free critical contacts, each at three distinct points.

References

1. M. Atallah, Dynamic computational geometry, *Proceedings of the 24th Symposium on Foundations of Computer Science*, 92-99, 1983.
2. S. Hart and M. Sharir, Nonlinearity of Davenport-Schinzel sequences and of generalized path compression schemes, *Combinatorica* **6** (1986), 175-201.
3. K. Kedem and M. Sharir, An Efficient Motion-Planning Algorithm for a Convex Polygonal Object in Two-Dimensional Polygonal Space, Technical Report 253, Computer Science Department, Courant Institute, 1986.
4. K. Kedem, R. Livne, J. Pach, and M. Sharir, On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles, *Discrete Comput. Geom.* **1** (1986), 59-71.
5. D. Leven and M. Sharir, An efficient and simple motion-planning algorithm for a ladder moving in two-dimensional space amidst polygonal barriers, *Proceedings of the ACM Symposium on Computational Geometry*, 221-227, 1985 (also to appear in *J. Algorithms*).
6. D. Leven and M. Sharir, Planning a Purely Translational Motion for a Convex Object in Two-Dimensional Space Using Generalized Voronoi Diagrams, Technical Report 34/85, The Eskenasy Institute of Computer Science, Tel Aviv University, 1985 (also to appear in *Discrete Comput. Geom.*).
7. 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, *Comm. Pure Appl. Math.* **36** (1983), 345-398.
8. M. Sharir, Almost Linear Upper Bounds on the Length of Generalized Davenport-Schinzel Sequences, Technical Report 29/85, The Eskenasy Institute of Computer Science, Tel-Aviv University, 1985 (also to appear in *Combinatorica* **7** (1987).)
9. M. Sharir, Improved Lower Bounds on the Length of Davenport-Schinzel Sequences, Technical Report 204, Computer Science Department, Courant Institute, 1986.
10. E. Szemeredi, On a problem by Davenport and Schinzel, *Acta Arith.* **25** (1974), 213-224.

Received October 2, 1985, and in revised form September 2, 1986.