TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 298, Number 1, November 1986

SIMPLEXWISE LINEAR UNTANGLING

DAVID W. HENDERSON

ABSTRACT. In this paper we show how to canonically untangle simplexwise linear spanning arcs of a convex 2-cell. Specifically, we show that the space of such arcs is contractible. The main step in the contraction is a flow along the gradient field of an energy function. A 3-dimensional version of this result would imply the Smale Conjecture—Hatcher Theorem.

0. Introduction. We will show how to canonically untangle simplexwise linear arcs in the plane. Specifically, let K^2 be a triangulated convex 2-cell in $\mathbf{R}^2$. Let v and w be any two points that lie in the interior of distinct 1-simplexes in the boundary of K^2. Let J_k be the (straight) segment joining v to w and subdivided into k subintervals. Let $\mathscr{T}(J_k, K^2)$ denote the space of all maps $f\colon J_k \to K^2$, which are simplexwise linear (i.e. linear on each 1-simplex of J_k) and are such that

$$f(v) = v, \quad f(w) = w, \quad \text{and} \quad f(J_k) \cap \partial K^2 = \{v\} \cup \{w\},$$

where ∂ denotes the boundary. Since $f \in \mathscr{T}(J_k, K^2)$ is determined by its values on the vertices $v_1, v_2, \ldots, v_{k-1}$ of I_k, we can identify $\mathscr{T}(J_k, I^2)$ with an (open) subset of $(\mathbf{R}^2)^{k-1}$, where f is identified with the $(k-1)$-tuple $(f(v_1), f(v_2), \ldots, f(v_{k-1}))$. We shall prove

THEOREM. *$\mathscr{T}(J_k, K^2)$ is contractible.*

This is related to the main result of **[BCH]** which proves that the space of simplexwise linear homeomorphisms of K^2, rel ∂K^2, is contractible. See **[BCH]** and **[CHHS]** for a discussion of history, results and questions concerning spaces of simplexwise maps. The main interest in the Theorem and its proof is that it gives a method for attempting a proof of the 3-dimensional version:

CONJECTURE. *If A is a polyhedral spanning arc of I^3 and $\mathscr{T}(A, I^3)$ is the space of simplexwise linear unknotted embeddings of A into I^3,* rel ∂I^3, *then $\mathscr{T}(A, I^3)$ is contractible.*

See §6 for a discussion of this conjecture. According to a result in **[Hen]**, this conjecture, if true, would imply the analogous result in the smooth category. This later result is known to be equivalent to the well-known Smale Conjecture, recently proved by A. Hatcher (see **[Hat2]**).

Received by the editors March 31, 1985 and, in revised form, November 4, 1985.

The author was partially supported by NSF Grant MCS 83-01865. Early stages of this research were done while the author was a visiting faculty member at Birzeit University, West Bank.

1980 *Mathematics Subject Classification*. Primary 58D10, 57Q35; Secondary 578M99.

0002-9947/86 $1.00 + $.25 per page

1. Reduction of the Theorem. If P is a projective transformation of the extended plane which maps at most one point of K to infinity and if $f \in \mathscr{J}(J_k, K^2)$, then define f_P: $P(J_k) \to P(K)$ by setting $f_P(z) = P \circ f \circ P^{-1}(z)$ for z a vertex of J_k, and extending linearly. Note that $f_P \in \mathscr{T}(P(J_k), P(K^2))$ and $f_P(P(J_k)) = P \circ f \circ P^{-1}(P(J_k))$, but $f_P \neq P \circ f \circ P^{-1}$, since P is not linear on simplexes. Now pick a projective transformation P so that $P(J_k) = [0, k] \times \{0\}$ and so that the P-images of the 1-simplexes containing the endpoints of J_k are vertical and contain the segments $\{0\} \times [-1, 1]$ and $\{k\} \times [-1, 1]$ ($\subset \mathbf{R}^1 \times \mathbf{R}^1$) (see Figure 1). Note that if v and w belong to adjacent edges of ∂K, then P will send the vertex between v and w to ∞.

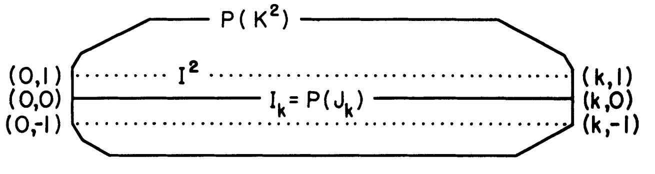

FIGURE 1

(1.1) LEMMA. *$\mathscr{T}(J_k, K^2)$ is homeomorphic to $\mathscr{T}(P(J_k), P(K^2))$ which contracts within itself into the subspace $\mathscr{T} \equiv \mathscr{T}(I_k, I^2)$, where $I_k \equiv P(J_k) = [0, k] \times \{0\}$ and $I^2 \equiv [0, k] \times [-1, 1]$.*

PROOF. It is easy to check that $f \to f_P$ is a homeomorphism with inverse $g \to g_{P^{-1}}$. The contraction is accomplished by composing each function in $\mathscr{T}(P(J_k), P(K^2))$ with the homotopy of $\mathbf{R}^1 \times \mathbf{R}^1$, which takes

$$(x, y) \mapsto (x, (1 - t + td)y)$$

at time t for $d = \max\{\|y\| \mid y$ is the y-coordinate of some point in $P(K^2)\}^{-1}$. □

In the remainder of the paper we will prove the Theorem by showing that $\mathscr{T} \equiv \mathscr{T}(I_k, I^2)$ is contractible.

2. The contraction of $\mathscr{T}$. In order to contract $\mathscr{T}$ it is sufficient to show that each compact subset $\mathscr{C} \subset \mathscr{T}$ can be contracted in $\mathscr{T}$. For $f \in \mathscr{C}$, we proceed by straightening out $f(I_k)$ one vertex at a time. That is, if

$$\mathscr{T}_j \equiv \{f \in \mathscr{T} \mid f(i) = i \text{ for } i = 1, 2, \ldots, j\},$$

then we homotope $\mathscr{C}$ into $\mathscr{T}_1$ and then homotope (in $\mathscr{T}_1$) the (compact) image of $\mathscr{C}$ into $\mathscr{T}_2$, and so forth, until we reach $\mathscr{T}_{k-1}$, which consists only of the identity map. The homotopy of a compact subset of $\mathscr{T}_j$ (in $\mathscr{T}_j$) is accomplished in two steps:

(2.1) LEMMA. *Any compact subset of $\mathscr{T}_j$ is homotopic (in $\mathscr{T}_j$) to a (compact) subset of*

$$\mathscr{T}_j^* \equiv \{f \in \mathscr{T}_j \mid \pi_x f(j + 1) = j + 1\},$$

where π_x is the projection of $\mathbf{R}^1 \times \mathbf{R}^1$ onto $\mathbf{R}^1 \times \{0\}$.

(2.2) LEMMA. *Any compact subset of $\mathscr{T}_j^*$ is homotopic (in $\mathscr{T}_j^*$) to a (compact) subset of $\mathscr{T}_{j+1}$.*

PROOF OF LEMMA (2.1) [AND LEMMA (2.2)]. The construction of the homotopy is very similar in both steps: In the main text the homotopy of (2.1) will be detailed [the changes necessary for (2.2) will be indicated in square brackets]. The homotopy is defined to be horizontal [vertical]; i.e. $\pi_y f_t(i)$ [$\pi_x f_t(i)$] will be constant. With these restrictions the homotopy is defined as the flow along the gradient vector field of an energy function $E\colon \mathscr{T}_j \to [0, \infty)$ [$E\colon \mathscr{T}_j^* \to [0, \infty)$]. That is, the homotopy $f \mapsto f_t$ is defined by letting f_t be the solution of the system of equations

$$\frac{d}{dt} f_t(i) = \frac{-\partial}{\partial x(i)} E(f_t) \quad \left[= \frac{-\partial}{\partial y(i)} E(f_t) \right], \qquad f_0 = f,$$

where $f_t(i) = (x(i), y(i))$. Part of E is a "spring energy", S, acting as if there were a spring pulling $f(j+1)$ to where it is supposed to go. That is, for (2.1), define S on $\mathscr{T}_j$ by setting $S(f) = \|\{j+1\} - \pi_x f(j+1)\|^2$. [For (2.2), define S on $\mathscr{T}_j^*$ by setting $S(f) = \|\{j+1\} - f(j+1)\|^2$.] If this were all of E, then in most cases during the flow one of the vertices $f_t(i)$ would intersect the boundary, ∂I^2, of I^2 or some $f_t([r, r+1])$ for $r \neq i-1, i$. To prevent this from happening, some "bumper energies", B, are added, enabling points and segments to be kept apart. In particular, let $h\colon (0, \infty) \to [0, \infty)$ be defined by

$$h(x) = \begin{cases} \dfrac{A}{x} \exp\left(\dfrac{-1}{d - x} \right) & \text{for } 0 < x < d, \\ 0 & \text{for } d \leqslant x, \end{cases}$$

which is a monotone decreasing C^∞ function depending on the two parameters A and d, which will be defined in §4. Then, if p is a point in the interior of I^2 and Q is ∂I^2 or a line segment disjoint from p, define $B(p, Q) \equiv h(D(p, Q))$, where $D(p, Q) \equiv \inf\{\|p - q\| \mid q \in Q\}$ is the distance from p to Q. Notice that $B(p, Q)$ is zero if p and Q are further apart than d and that, as p approaches Q, $B(p, Q) \to \infty$. We can now define

$$E(f) \equiv S(f) + \sum B(p, Q),$$

where the sum is over all $p = f(i)$ and $Q = \{\partial I^2 \text{ or } f([r, r+1]) \text{ for } r \neq i-1, i\}$. For $f \in \mathscr{T}_j$, the first j vertices of I_k are fixed and thus we can represent $f = (f(j+1), f(j+2), \ldots, f(k+1)) \in (\mathbf{R}^2)^{k-1-j}$ and $\mathscr{T}_j$ is an open subset of $(\mathbf{R}^2)^{k-1-j}$ and $\mathscr{T}_j^*$ is an open subset of $(\{j+1\} \times \mathbf{R}) \times (\mathbf{R}^2)^{k-2-j}$. Define f_t to be the unique solution of the system of differential equations

$$\frac{d}{dt} f_t(i) = \frac{-\partial}{\partial x(i)} E(f_t), \quad f_0 = f \quad \left[\frac{d}{dt} f_t(i) = \frac{-\partial}{\partial y(i)} E(f_t),\ f_0 = f \right]$$

for $j+1 \leqslant i \leqslant k-1$. The theory of ordinary differential equations applies because in §3 we will show that $\partial E/\partial x(i)$ [$\partial E/\partial y(i)$] are Lipschitz functions of $f \in \mathscr{T}_j$ [$\mathscr{T}_j^*$]. In order for f_t to move outside of $\mathscr{T}_j$ [$\mathscr{T}_j^*$] it is necessary that either f_t fails to be one-to-one or $f_t(i)$ intersects ∂I^2 for some $i = j+1, j+2, \ldots, k-1$, but in either case some $B(f_t(i), Q)$ will be infinite, which is impossible since the energy $E(f_t)$ starts finite and, for $z = x$ [$z = y$],

$$\frac{d}{dt} E(f_t) = \nabla E(f_t) \cdot \frac{d}{dt} f_t = -\sum_{i=j+1}^{k-1} \left\| \frac{\partial}{\partial z(i)} E(f_t) \right\|^2 < 0.$$

Thus, if f is in the compact $\mathscr{C} \subset \mathscr{T}_j$ $[\mathscr{T}_j^*]$, then f_t stays in the compact $E^{-1}E(\mathscr{C}) \subset \mathscr{T}_j$ $[\mathscr{T}_j^*]$. It follows from the standard theory of differential equations (see, for example, [**Spi**, pp. 194–200] that $(f, t) \to f_t$ defines a continuous flow on $\mathscr{T}_j$ $[\mathscr{T}_j^*]$ for $t \in [0, \infty)$. Let

$$\nabla_z E \equiv \left(\frac{\partial}{\partial z(j+1)} E, \ldots, \frac{\partial}{\partial z(k-1)} \right) \quad \text{for } z = x \; [z = y].$$

In §5 we will show

(2.3) LEMMA. *$S(f) \neq 0$ implies that $\nabla_z E(f) \neq 0$ for $z = x$ $[z = y]$.*

Then we conclude that, for all $S(f_t) > \varepsilon$, there is a δ such that

$$\frac{d}{dt} E(f_t) = \nabla_z E(f_t) \cdot \frac{d}{dt} f_t = -\| \nabla_z E(f_t) \|^2 < \delta < 0,$$

and thus $\lim_{t \to \infty} S(f_t) = 0$, since $E(f_t)$ is bounded below by 0. Now, we do not know that this flow can be extended to $t = \infty$; however, eventually $f_t(j+1)$ will be so close to its destination that we can stop the flow (for all $j + 1 \leqslant i \leqslant k - 1$ simultaneously) and continue the homotopy by moving only $f(j+1)$ along the straight line to its destination, $(j+1, \pi_y f(j+1))$ $[(j+1, 0)]$. This finishes the proof of (2.1) [(2.2)]. □

3. $\partial E/\partial x(i)$ and $\partial E/\partial x(i)$ are Lipschitz functions. Since

$$\frac{\partial}{\partial z} E(f) = \frac{\partial}{\partial z} S(f) + \sum h'(D(p,Q)) \frac{\partial}{\partial z} D(p,Q)$$

for $z = x(i)$ $[y(i)]$, it suffices to show that $(\partial/\partial z) D(p,Q)$ is Lipschitz for $p = f(s)$ and $Q = \{\partial I^2$ or $f([r, r+1])$, $r \neq s-1, s\}$. We see that if $p = f(s)$ and $Q = f([r, r+1])$, then

$$\nabla_i D \equiv \left(\frac{\partial}{\partial x(i)} D(p,Q), \frac{\partial}{\partial x(i)} D(p,Q) \right)$$

is nonzero only if $i = s$, r or $r + 1$. We use the fact that the gradient, $\nabla_i D$, is in the direction that $f(i)$ must be moved to obtain the maximum rate of increase of D and that that rate is the magnitude of $\nabla_i D$. Let q be the closest point of Q to p. We have three cases: $q = f(r)$, $q = f(r+1)$, or $q \in f((r, r+1))$.

(i) $q = f(r)$. Then $D(p,Q) = \|f(s) - f(r)\|$ and

$$\nabla_s D = \frac{f(s) - f(r)}{\|f(s) - f(r)\|} = -\nabla_r D \quad \text{and} \quad \nabla_{r+1} D = 0.$$

(ii) $q = f(r+1)$. Then $D(p,Q) = \|f(s) - f(r+1)\|$ and

$$\nabla_s D = \frac{f(s) - f(r+1)}{\|f(s) - f(r+1)\|} = -\nabla_{r+1} D \quad \text{and} \quad \nabla_r D = 0.$$

(iii) $q \in f((r, r+1))$. Then, if μ is the unit vector perpendicular to Q and pointing towards p, we have

$$\nabla_s D = \mu, \quad \nabla_r D = \frac{-\|q - f(r+1)\|}{\|f(r) - f(r+1)\|} \cdot \mu, \quad \text{and}$$

$$\nabla_{r+1} D = \frac{-\|q - f(r)\|}{\|f(r) - f(r+1)\|} \cdot \mu.$$

We notice that these gradients are continuous, but not differentiable. But restricted to each of the three pieces each gradient is differentiable and bounded (in $\mathscr{T}$). Thus these gradients are Lipschitz.

4. The constants c, A, d, **and the orderings** $<$ **and** o. Throughout this section $\mathscr{C}$ is some compact subset of $\mathscr{T}_j$ $[\mathscr{T}_j^*]$. Notice that the homotopy of $\mathscr{T}_j$ $[\mathscr{T}_j^*]$ into $\mathscr{T}_j^*$ $[\mathscr{T}_{j+1}]$ will be contained in the set $\mathscr{C}^{\sim}$, $\mathscr{C} \subset \mathscr{C}^{\sim} \subset \mathscr{T}_j$ $[\mathscr{T}_j^*]$, where we set $g \in \mathscr{C}^{\sim}$ iff there is an $f \in \mathscr{C}$ such that

(a) $\pi_y f(t) = \pi_y g(t)$ $[\pi_x]$ for all $t \in I_k$.

(b) for all $t, t' \in I_k$, $\pi_y f(t) = \pi_y(t')$ $[\pi_x]$ implies that $\{\pi_x f(t), \pi_x f(t')\}$ has the same order as $\{\pi_x g(t), \pi_x g(t')\}$ $[\pi_y]$.

(c) $E(g) \leqslant E(\mathscr{C}) \equiv \mathrm{Max}\{E(f') \mid f' \in \mathscr{C}\}$.

(4.1) LEMMA. *There is a constant* $c \equiv c(\mathscr{C}) < \frac{1}{4}$ *such that, for all* $1 \leqslant r, s \leqslant k - 1$ *and* $g \in \mathscr{C}^{\sim}$, *whenever the horizontal [vertical] line through* $g(s)$ *intersects* $g([r-1, r])$ *and* $g([r, r+1])$ *on opposite sides of* $g(s)$, *then* $\|\pi_y g(s) - \pi_y g(r)\| \geqslant c$ $[\|\pi_x g(s) - \pi_x g(r)\| \geqslant c]$.

PROOF. We shall prove (4.1) only for the case of $\mathscr{T}_j$ and leave to the reader the modifications necessary for $\mathscr{T}_j^*$. Suppose that $\{g_i\} \subset \mathscr{C}^{\sim}$ such that the horizontal line through $g_i(s)$ intersects $g_i([r-1, r])$ at p_i and intersects $g_i([r, r+1])$ at q_i on opposite sides of $g(s)$ and suppose that

$$\|\pi_y g_i(s) - \pi_y g_i(r)\| \to 0.$$

Then, for the corresponding $f_i \in \mathscr{C}$, some subsequence of $\{f_i\}$ will converge to $f \in \mathscr{C}$, since $\mathscr{C}$ is compact. Thus

$$\lim \pi_y(p_i) = \lim \pi_y(q_i) = \pi_y f(s) = \pi_y f(r)$$

and $\pi_x f(s)$ is between $\lim \pi_x(p_i)$ and $\lim \pi_x(q_i)$. Therefore, $f(s) \in f([r-1, r+1])$, which is impossible since $f \in \mathscr{T}$ is one-to-one. □

(4.2) DEFINITIONS. (i) Set $b \equiv b(\mathscr{C}) \equiv 2(k/c)^2 > 2k/c > 8k$.

(ii) Choose $d = d(\mathscr{C}) > 0$ so small that $b^{2k}d \leqslant c/2$.

(iv) Pick $A = A(\mathscr{C})$ so large that $h(x) \leqslant E(\mathscr{C})$ implies $\frac{3}{4}d \leqslant x$.

For what follows, assume that $f \in \mathscr{C}^{\sim}$. We now use the above defined constants to construct several orderings on the segments: $P_r = f([r, r+1])$, $0 \leqslant r \leqslant k-1$.

(4.3) DEFINITION. (a) $P_i <_{\mathbf{a}} P_r$ iff there is a horizontal [vertical] line in I^2 which first (traveling from left to right [down to up]) intersects P_i and then intersects P_r.

(b) $P_i <_{\mathbf{b}} P_r$ iff $P_i <_{\mathbf{a}} P_{r_1} <_{\mathbf{a}} P_{r_2} <_{\mathbf{a}} \cdots <_{\mathbf{a}} P_{r_n} <_{\mathbf{a}} P_r$, for some n.

(c) $P_i <_{\mathbf{c}} P_r$ iff P_r not-$<_{\mathbf{b}} P_i$ and the projections $\pi_y P_i$ $[\pi_x P_i]$ and $\pi_y P_r$ $[\pi_x P_r]$ are disjoint and, for all those endpoints, $p \in P_i$ and $q \in P_r$, whose π_y $[\pi_x]$-images are closest, it is true that these images are within d of each other (see (4.2)), and $\pi_x p < \pi_x q$ $[\pi_y]$ or $\{\pi_y(\pi_y \mid P_i)^{-1}\pi_y p \cap \pi_x(\pi_y \mid P_r)^{-1}\pi_y q$ is nonempty and $\pi_y p < \pi_y q\}$ [interchange π_x and π_y].

(d) $P_i < P_r$ iff $P_i <_{\mathbf{e}_0} P_{r_1} <_{\mathbf{e}_1} P_{r_2} <_{\mathbf{e}_2} \cdots <_{\mathbf{e}_{n-1}} P_{r_n} <_{\mathbf{e}_n} P_r$ for some n, where for each i, $\mathbf{e}_i = \mathbf{a}$ or $\mathbf{e}_i = \mathbf{c}$.

Notice that $<_{\mathbf{c}}$ is discontinuous with respect to f in the sense that, for some f, $P_i <_{\mathbf{c}} P_r$ but, for some g arbitrarily near f, $P_r <_{\mathbf{c}} P_i$. This is the main source of the g_f constructed in §5 being discontinuous if f. We now prove

(4.4) LEMMA. *It is not true that $P_r < P_r$. Thus $<$ is a partial ordering on the $\{P_i\}$ for each $f \in \mathscr{C}^{\sim}$.*

PROOF. Suppose $P_{r_1} < P_{r_2} < \cdots < P_{r_n}$ are adjacent elements in the ordering, i.e., for each i, there is no P_s such that $P_{r_i} \leqslant P_s \leqslant P_{r_{i-1}}$. It follows that, for each $1 \leqslant i \leqslant n-1$, $P_{r_i} \leqslant_{\mathbf{e}_i} P_{r_{i+1}}$, where $\mathbf{e}_i = \mathbf{a}$ or $\mathbf{e}_i = \mathbf{c}$. Let $p_i \in P_{r_i}$ and $q_{i+1} \in P_{r_{i+1}}$ be such that $\pi_y(p_i)\pi_y(q_{i+1})$ $[\pi_y]$ if $\mathbf{e}_i = \mathbf{a}$, or $\|\pi_y(p_i - q_{i+1})\| < a$ $[\pi_x]$ if $\mathbf{e}_i = \mathbf{c}$. Look at the piecewise-linear path

$$\Lambda \equiv p_1 - q_2 - p_2 - q_3 - p_3 - \cdots - q_{n-1} - p_{n-1} - q_n.$$

At each vertex of Λ whose y $[x]$-coordinate is a local extremum the path turns towards the right [top] and thus the first (in traversing Λ from p_1 to q_n) point $p \in \Lambda$, if any, which intersects a point on the path before p must be as in Figure 4.

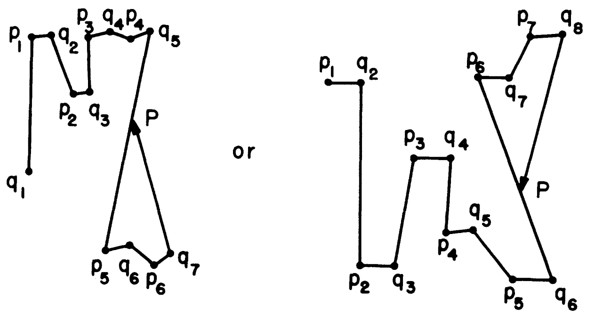

FIGURE 4

If there is such a p then there are four possibilities: For some $r < s$, (i) $p = [q_r, p_r] \cap [q_s, p_s]$, (ii) $p = [q_r, p_r] \cap [p_s, q_{s+1}]$, (iii) $p = [p_r, q_{r+1}] \cap [q_s, p_s]$, or (iv) $p = [p_r, q_{r+1}] \cap [p_s, q_{s+1}]$. Since $[q_i, p_i] \subset P_{r_i}$, if (i) were true then we would have $P_{r_r} = P_{r_s} \equiv P$ and from Figure 4 we see that p_r, q_r must be in the same order on P as q_s, p_s and thus either $p_r \in [q_s, p_s]$ or $q_s \in [p_r, q_r]$. However, both of these are impossible since q_{r+1} is left of [below] p_r and p_{s-1} is right of [above] q_s and p is the first point of intersection; thus we conclude that (i) is impossible. If $\mathbf{e}_s = \mathbf{a}$ then (ii) is impossible since p_s and p_{s+1} are assumed to be adjacent in the ordering. If $\mathbf{e}_s = \mathbf{c}$ then (ii) would imply that P_{r_r} is less than $\frac{3}{4}d$ from either P_{r_s} or $P_{r_{s+1}}$, which is not possible by (4.2). Similarly, (iii) is impossible. From Figure 4 we see that at least one of the intervals of Λ containing p must point towards the left [bottom], thus (iv) is impossible. □

In the construction of g_f in §5, we shall use the above partial ordering of the P_i and a related total ordering induced by a bijection:

(4.5) DEFINITION. Choose o: $\{0, 1, \ldots, k-1\} \to \{0, 1, \ldots, k-1\}$ to be a one-to-one function so that:

(a) If $o(r) < o(s)$, then either $P_r < P_s$, or P_r and P_s are incomparable.

(b) If no segment $P_i \in f(I_k)$ is between (in the $<$ -partial ordering) two adjacent segments $\{P_{r-1}, P_r\}$, then $o(r-1) = o(r) + 1$, or $o(r) = o(r+1) + 1$.

(c) If $S \equiv \{P_0, P_1, \ldots, P_s\}$ is maximal with respect to satisfying (b) for each pair $\{P_r, P_{r+1}\}$, $0 \leqslant r < s$, then $o(i) < o(0)$ iff $P_i < P_r$ for some $r \leqslant s$. (Note that $j \leqslant s$.)

Define $m(i) \equiv \min\{o(i), o(i-1)\}$ and $M(i) \equiv \operatorname{Max}\{o(i), o(i-1)\}$. Note that $f(i)$ belongs to P_{i-1} and P_i, and so $m(i)$ gives the first place and $M(i)$ the last place that $f(i)$ is represented in the ordering o.

(4.6) LEMMA. *Let $f \in \mathscr{C}^{\sim}$. If $\pi_y f(i)$ is within d of $\pi_y(P_{r-1} \cup P_r)$ $[\pi_x]$ and $\pi_y f(i)$ $[\pi_x]$ is within $c = c(\mathscr{C})$ of $\pi_y f(r)$ $[\pi_x]$, then either $m(i) \geqslant M(r)$ or $M(i) \leqslant m(r)$, with equality only if $i = r \pm 1$.*

PROOF. By (4.1) it is impossible that $f(i)$ be between P_r and P_{r-1}. Therefore, either (a) $\pi_y f(i)$ intersects both $\pi_y P_r$ and $\pi_y P_{r-1}$ $[\pi_x]$, (b) $\pi_y P_r \cap \pi_y P_{r-1} = \pi_y f(r)$ $[\pi_x]$, or (c) π_{P_r} or $\pi_y P_{r-1}$ is contained in $[\pi_y f(i), \pi_y f(r)]$ $[\pi_x]$. Thus, by (4.1) and (4.5)(b), we see in each case that neither P_{i-1} nor P_i can come between P_r and P_{r-1} in the $>$ -ordering. Therefore the desired inequalities hold with equality only if $P_r = P_{i-1}$ or $P_{r-1} = P_i$. $\square$

5. Proof of Lemma (2.3). Preparatory to proving (2.3) at the end of this section, we construct a certain function $g_f \in \mathscr{C}^{\sim}$ for each $f \in \mathscr{C}^{\sim}$.

(5.1) LEMMA. *If $f \in \mathscr{C}^{\sim}$ and $S(f) \neq 0$, then there is a $g_f \in \mathscr{C}^{\sim}$ such that*:

(a) $\pi_y g_f(i) = \pi_y f(i)\,[\pi_x]$, $0 \leqslant i \leqslant k$,

(b) *for all* $0 \leqslant s \leqslant 1$, $sf + (1-s)g_f \in \mathscr{T}_j$ $[\mathscr{T}_j^*]$,

(c) $E(g_f) = S(g_f) = 0$,

(d) $(f(j+1) - g_f(j+1)) \cdot \nabla_{j+1} S(f) > 0$,

(e) *for all* $0 \leqslant s \leqslant 1$, $B(sf + (1-s)g_f) < B(f)$, *where* $B(\cdot)$ *ranges over all* $B = B(p, Q)$ *that appear in the definition of* $E(\cdot)$.

We will prove (5.1) after giving the inductive construction of g_f in (5.2). The g_f which we construct will not vary continuously with f; in fact, note that if we could find the g_f that varied continuously with f, then (5.1) would produce the desired homotopy of $\mathscr{T}_j$ $[\mathscr{T}_j^*]$.

In order to construct g_f we make use of the constants d, b, c, and the orderings $<$, o of §4. Recall $m(i) \equiv \min\{o(i), o(i-1)\}$ and $M(i) \equiv \operatorname{Max}\{o(i), o(i-1)\}$, and let $r \equiv o^{-1}$: $\{0, 1, \ldots, k-1\} \to \{0, 1, \ldots, k-1\}$. We define g_f inductively, one segment at a time. In the case that $\pi_x f(j+1) > j+1$ $[\pi_y f(j+1) > 0]$, we define g_f first on $[r(0), r(0)+1]$ and then on $[r(1), r(1)+1]$ and so on in order until $[r(k-1), r(k-1)+1]$. In the case that $\pi_x f(j+1) < j+1$ $[\pi_y f(j+1) < 0]$, we define g_f on the segments in the reverse order. We shall describe in detail only the first case; the changes necessary for the second case should be clear to the reader.

$(5.2)_s$ LEMMA. *In the case that* $\pi_x f(j+1) > j+1$ $[\pi_y(j+1) > 0]$, *the function* g_f *can be defined on*

$$Cs \equiv \bigcup_{i=0}^{s-1} [r(i), r(i)+1]$$

so that the following hold:

$(a)_s$ g_f *is linear on each interval* $[i, i+1] \in Cs$, $\pi_y g_f(p) = \pi_y f(p)$ $[\pi_x g_f(p) = \pi_x f(p)]$, *and, for* $0 \leqslant i \leqslant j$, $g_f(i) = (i, 0)$ *and* $g_f(k+1) = (k+1, 0)$, *whenever defined.*

$(b)_s$ *There is a finite subset*

$$Ts \equiv \{ g_f(i) | M(i) \geqslant s \} \cup \{ p_1, p_2 \} \subset I^2,$$

where $p_1 = (b^{2k}d, \operatorname{sign}(\pi_y f(j+1))1)$ $[= (0, -1 + b^{2k}d)]$ *and* $p_2 = (b^{2k}d, 0)$ $[= (k, -1 + b^{2k}d)]$ *if* $s \leqslant o(0)$, *and* $p_1 = (j+1+b^{2k}d, -)$ $[= (0, b^{2k}d)]$ *and* $p_2 = (j+1+b^{2k}d, 1)$ $[= (k, b^{2k}d)]$ *if* $s > o(0)$. *For* $p \in Ts$, *define* $M(p) = M(i)$ *if* $p = g_f(i)$, *and* $M(p_1) = M(p_2) = o(0)$ *if* $s \leqslant o(0)$, *and* $M(p_1) = M(p_2) = k$ *if* $s > o(0)$. *We order Ts by the order on* $\pi_y Ts$ $[\pi_x Ts]$. *Except for* p_1, p_2, *Ts is just those vertices of* $g_f(Cs)$ *which are endpoints of only one simplex in* $g_f(Cs)$.

$(c)_s$ *Let* q, q' *be in Ts. If, for every r between q and q' in Ts,* $M(r) < M(q)$ *and* $M(r) < M(q')$ *and if* $[q, q']$ *is not contained in Cs, then, for all* $0 < \tau < 1$, *either* (i) $p \equiv \tau \cdot q + (1-\tau) \cdot q'$ *is at least as far horizontally* [*vertically*] *as*

$$D(q, q'; \tau; s) \equiv \left(\tau b^{2M(q)} + (1-\tau) b^{2M(q')}\right) d(b^{-1} - sb^{-2})$$
$$= \tau\left(b^{2M(q)-1} - sb^{2M(q)-2}\right)d + (1-\tau)\left(b^{2M(q)-1} - sb^{2M(q)-2}\right)d$$

from any point within d of Cs or (ii) $\pi_y p$ $[\pi_x p]$ *is within c of* $\pi_y q$ (*or* $\pi_y q'$) $[\pi_x]$ *and p is closer horizontally* [*vertically*] *than* $D(q, q'; \tau; s)$ *to the* g_f*-image of some interval containaing q* (*or* q'), *where* $M(p) \equiv \tau \cdot M(q) + (1-\tau) \cdot M(q)$ (*see Figure* 5.1).

$(d)_s$ *Let* $q \in Ts$. *If* $M(q) < o(0)$ *and* $p \in Jq$ *with* $\pi_y p = \pi_y f(j+1)$ $[\pi_x]$, *then* $(j+1) - \pi_x p$ $[0 - \pi_y p] \geqslant b^{2k}d$. *If* $M(q) > o(0)$ *and* $q \in Ts$, *then* (c) *holds for* $q' \equiv (j+1+b^{2k}d, \pm 1)$ $[\equiv (0, b^{2k}d)$ *or* $(k, b^{2k}d)]$ *with* $M(q') \equiv k$.

$(e)_s$ *If* $m(i) \leqslant o(0) \leqslant M(i)$, *then* $j+1 \leqslant \pi_x g_f(i) \leqslant j+1+b^{2k}d$ $[0 \leqslant \pi_y g_f(i) \leqslant b^{2k}d]$.

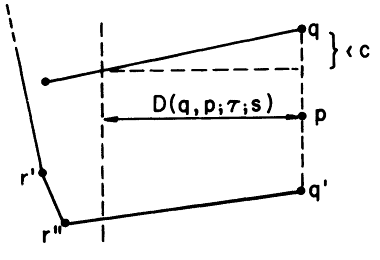

FIGURE 5.1

PROOF THAT $(5.2)_s$ IMPLIES $(5.2)_{s+1}$. We add $[r(s), r(s)+1]$ to Cs and have three cases depending on whether both, neither, or only one of $r(s)$, $r(s)+1$ are in Cs.

Both $r(s)$ *and* $r(s)+1$ *are in* Cs. In this case $q = q_f(r(s))$ and $q' = g_f(r(s)+1)$ are in Ts. If $\pi_y g_f(i)$ is between $\pi_y q$ $[\pi_x q]$ and $\pi_y q'$ $[\pi_x q']$, then $M(i) < s = M(s)$ and therefore $q_f(i) \notin Ts$. Thus q and q' are adjacent (in the π_y $[\pi_x]$-induced order on Ts) and, by $(c)_s$, we can extend g_f linearly on $[r(s), r(s)+1]$ and it will still be one-to-one and will satisfy $(a)_{s+1}$. The two points on either side of q and q' in Ts are now adjacent in $Ts+1$. The other parts of $(5.2)_{s+1}$ now follow if you note that, for each $q \in Ts$, $s < M(q)$ and thus that $b^{2s} \leqslant b^{2M(q)-2}$ (see Figure 5.2).

Only one of $r(s)$ *and* $r(s)+1$ *are in* Cs. Say $r(s)$ is not in Cs and $r(s)+1$ is in Cs (and $q = g_f(r(s)+1) \in Ts$). By the same argument as above, no $r \in Ts$ can be between q and $l \equiv \pi_y^{-1}\pi_y(r(s))$ $[\equiv \pi_s^{-1}\pi_x(r(s))]$ in the ordering induced by π_y $[\pi_x]$. Thus l must be between q and some $p \in Ts$ which is adjacent to q in Ts. Note that $M(q) = s$ and so all other $p \in Ts$ have larger M and thus will be in $Ts+1$. Let q_1 and q_2 be the first points in Ts on each side of l such that $M(q_1), M(q_2) > M(r(s))$, if there are any such. If on a side of l there is no such point q_1, then note that $M(r(s)) \geqslant M(p_i) \equiv o(0)$ and set

$$q_1 = \left(j + 1 + b^{2k}d, \pm 1\right) \quad \left[= (0, b^{2k}d) \text{ or } (k, b^{2k}d)\right].$$

If $\pi_y l = \tau\pi_y q_1 + (1-\tau)\pi_y q_2$ $[\pi_x]$, then set $q'' \equiv \tau q_1 + (1-\tau)q_2$ and define

$$\pi_x g_f\left(r(s)\right) = \pi_x q'' - D(q_1, q_2; \tau; s) + b^{2M(r(s))}d \quad [\pi_y]$$

Now define $q' \equiv g_f(r(s))$ and thus $M(q') = M(r(s))$. Extend linearly to $[r(s), r(s)+1]$ (see Figure 5.3).

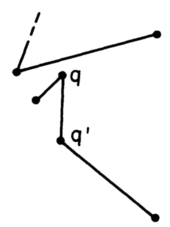

FIGURE 5.2

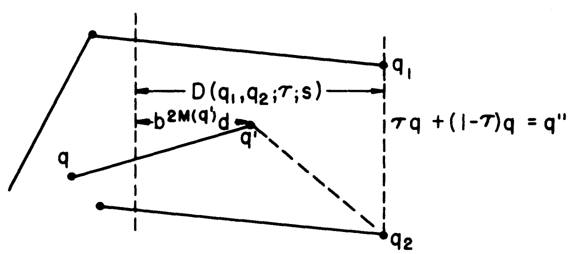

FIGURE 5.3

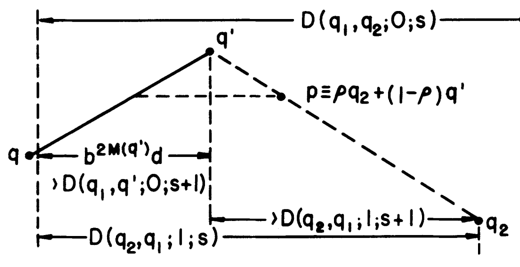

FIGURE 5.4

If $(c)_s$(ii) holds for $q_1, q_2; \tau$, then by (4.6) we conclude $M(q') \leqslant m(q_1)$ $\{$or $M(q') \leqslant m(q_2)\}$ which is impossible; thus we conclude that q'' is at least as far horizontally [vertically] from Cs as $D(q_1, q_2; \tau; s)$. Thus q' misses Cs and (see Figure 5.3)

$$\begin{aligned}\|q' - q''\| &= D(q_1, q_2; \tau; s) - b^{2M(q')}d \\ &= \{\tau b^{2M(q_1)} + (1-\tau)b^{2M(q_2)}\}(b^{-1} - sb^{-2})d - b^{2M(q')}d \\ &> D(q_1, q_2; \tau; s+1),\end{aligned}$$

since $M(q') \leqslant M(q_i) - 1$.

We must now check $(c)_{s+1}$. Let $q^{\sim} \equiv \alpha q_1 + (1-\alpha)q_2$ be such that $\pi_y(q^{\sim}) = \pi_y(q)$. Then, by $(c)_s$, $\|q - q^{\sim}\| > D(q_1, q_2; \alpha; s)$. Thus, for $\beta q' + (1-\beta)q \in g_f[r(s), r(s)+1]$, we have

$$\begin{aligned}\|\beta q' + (1-\beta)q - (\beta q'' + (1-\beta)q^{\sim})\| &= \beta\|q' - q''\| + (1-\beta)\|q - q^{\sim}\| \\ &\geqslant \beta D(q_1, q_2; \tau; s+1) + (1-\beta)D(q_1, q_2; \alpha; s+1) \\ &\geqslant D(q_1, q_2; \beta\tau + \alpha - \beta\alpha; s+1).\end{aligned}$$

For $\lambda \in [0,1] - [\alpha, \tau]$, the horizontally [vertically] closest point to $\lambda q_1 + (1-\lambda)q_2$ is in Cs. Thus $(c)_{s+1}$ holds for q_1, q_2.

All other cases of $(c)_{s+1}$ follow from $(c)_s$ except for q', q^*, where $\pi_y q^*$ is between or equal to $\pi_y q_1$ and $\pi_y q_2$. (Possibly q^* is q_1 or q_2.) We first set $q^* = q_2$ and look at $p = \rho q' + (1-\rho)q_2$, $0 \leqslant \rho \leqslant 1$ (see Figure 5.4). If the horizontal [vertical] distance from p to $[q, q']$ is less than $D(q', q_2; \rho; s+1)$, then, using similar triangles,

$$\begin{aligned}\|\pi_y p - \pi_y q'\| &< \{D(q_2, q'; \rho; s+1) - \rho D(q_2, q'; 1; s+1)\} \cdot \frac{\|\pi_y q - \pi_y q'\|}{b^{2M(q')}d} \\ &< (1-\rho)b^{2M(q')}d(b^{-1} - (s+1)b^{-2})\frac{k}{b^{2M(q')}d} < \frac{k}{b} < \frac{c}{2}\end{aligned}$$

(see Figure 5.4). It is also clear from Figure 5.4 that, if $\pi_y p$ is farther than c from $\pi_y q'$, then p is at least as far horizontally [vertically] as $D(q_2, q'; \rho; s+1)$ from any point which is within d of $Cs + 1$.

If remains now to look at $(c)_{s+1}$ with q' and $q^* \neq q_1, q_2$. Suppose that $\pi_y q^*$ is between $\pi_y q'$ and $\pi_y q_2$ and let $q^{\wedge}$ be the first point in Ts on the other side of q' from q^* such that $M(q^{\wedge}) > M(q^*)$. Notice that $q^*, q^{\wedge}$ satisfy the hypotheses of

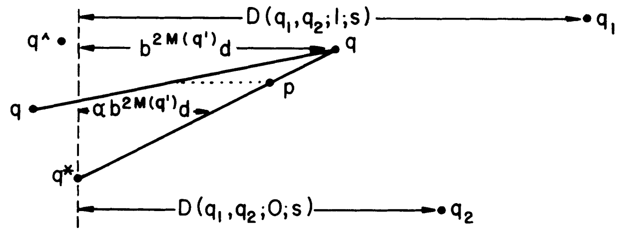

FIGURE 5.5

$(c)_s$. If $q^\wedge = q_1$, then an argument the same as that for q', q_2 will show that $(c)_{s+1}$ is satisfied for q', q^*. If $q^\wedge \neq q_1$, then the situation is as in Figure 5.5. Note that by Lemma (4.6), $\|\pi_y q - \pi_y q^*\| > c$ and thus that

$$\alpha \equiv \frac{\|\pi_y q - \pi_y q^*\|}{\|\pi_y q' - \pi_y q^*\|} > \frac{c}{k}.$$

If $p \equiv \beta q^* + (1 - \beta)q'$ is horizontally [vertically] within $D(q^*, q'; \beta; s + 1)$, then (since $M(q^*) < M(q')$)

$$\left\| \pi_y q' - \pi_y p \right\| \leqslant D\left(q^*, q'; \beta; s + 1\right) \cdot \frac{\|\pi_y q - \pi'_q\|}{\alpha b^{2M(q')} d}$$

$$< \left\{ \beta b^{2M(q^*) - 2M(q')} + (1 - \beta) \right\} \cdot \frac{k}{b} \cdot \frac{k}{c} < \frac{k^2}{bc} = \frac{c}{2}.$$

An examination of Figure 5.5 shows that $(c)_{s+1}$ holds for q', q^*.

Note that $(d)_{s+1}$ holds if $M(q') < o(0)$ because then q' is further from $g_f(j + 1)$ than the line joining q_1 and q_2, and neither q_1 nor q_2 is $(j + 1 + b^{2k}d, \pm 1)$ $[(0, b^{2k}d)$ or $(k, b^{2k}d)]$. If $M(q')) > o(0)$ and there is a $w \in Ts$ on the side of $\pi_y^{-1}\pi_y f(j + 1)$ $[\pi_x]$ with $M(w) \geqslant o(0)$, then $\pi_y[q, q']$ $[\pi_x[q, q']]$ does not contain $\pi_y f(j + 1)$ $[\pi_x]$. If $M(q') > o(0)$ and there is a $w \in Ts$, $M(w) \geqslant o(0)$, only on the side of q' away from $\pi_y^{-1}\pi_y f(($see$j + 1)$ $[\pi_x]$, then $w, p_2, p = [w, p_2] \cap \pi_y^{-1}\pi_y f(j + 1)$ $[\pi_x]$ satisfy $(c)_s$ and $(d)_s$ and, since $[q, q']$ must cross $[w, p_2]$, we see that $[q', p_2]$ must miss $Cs + 1$ and have slope at least $j[1 - b^{2k}d]$; but $|\pi_y f(j + 1) - \pi_y q'| \geqslant c$ $[\pi_x]$, and thus

$$j + 1 - \pi_x\left\{[q, q'] \cap \pi_y^{-1}\pi_y f(j + 1)\right\} \geqslant cj \geqslant b^{2k}d$$

$$\left[0 - \pi_y\left\{[q, q'] \cap \pi_x^{-1}\pi_x f(j + 1)\right\} \geqslant c(1 - b^{2k}d) \geqslant b^{2k}d\right]$$

(see (4.2) and Figure 5.6). If there is no $w \in Ts$, $M(w) \geqslant o(0)$, then $\pi_x g_f(Cs) < b^{2o(0)}$ $[\pi_y]$ and thus $[q, p_2]$ is as in the last sentence and we conclude that $(c)_{s+1}$ and $(d)_{s+1}$ hold.

Neither $r(s)$ *nor* $r(s) + 1$ *are in* Cs. The proof of this case is essentially the same as the previous case, except that we must pick two pairs of q_1, q_2, one for each of $r(s)$ and $r(s) + 1$. □

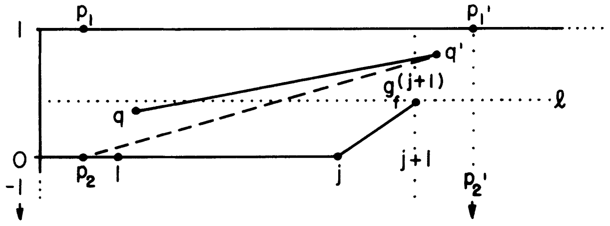

FIGURE 5.6

PROOF OF (5.1). The order < on the segments P_r (see (4.3)) ensures that if $B(f(i), P_r) \neq 0$, then $f(i)$ and P_r stay in the same order as $g_f(i)$ and $g_f([r, r+1])$. Since the homotopy is horizontal [vertical] and $D(g_f(i), g_f([r, r+1])) > d$, it follows that $B(sf + (1-s)g_f)$ must decrease. (5.1) now follows.

PROOF OF (2.3). Let $\mu \equiv g_f - f \subset (\mathbf{R}^2)^{k-1}$. Then from (5.1) we see that $\mu \cdot \nabla_y B(f) \leqslant 0$ $[\nabla_x]$ and $\mu \cdot \nabla_y S(f) < 0$ $[\nabla_x]$. Thus, $\mu \cdot \nabla_y E(f) < 0$ $[\nabla_x]$ and we conclude that $\nabla_y E(f) \neq 0$ $[\nabla_x]$. □

6. Comments on 3-dimensional conjecture. Let A be a straight unknotted spanning arc of I^3 which has a fixed subdivision. Let $\mathscr{T}^3$ be the space of simplexwise linear unknotted imbeddings of A into I^3, rel ∂I^3. It has been thought (see **[Hat1]**) that a proof that $\mathscr{T}^3$ is contractible is not feasible since there seems to be no way to use the hypothesis of unknottedness in a canonical way. However, it might be possible to define energy functions on $\mathscr{T}^3$ and attempt to flow along the gradient vector field in a manner similar to that above in the 2-dimensional case. The unknotted hypothesis would come in when one attempted to prove the analogue of Lemma (2.3), but in this case one would only be looking at a single $f \in \mathscr{T}^3$ and one could assume that the image of f bounds a disk in I^3 and not need to pick the disc canonically. This is analogous to the situation in §5 where the g_f did not depend continuously on f.

REFERENCES

[BCH] E. Bloch, R. Connelly and D. Henderson, *The space of simplexwise linear homeomorphisms of a convex 2-disk*, Topology **23** (1984), 161–175.

[CHHS] R. Connelly, D. Henderson, C.-w. Ho and M. Starbird, *On the problems related to linear homeomorphisms, embeddings, and isotopies*, Continua, Decompositions, Manifolds, Univ. of Texas Press, Austin, Texas, 1983, pp. 229–239.

[Hat1] A. Hatcher, *Linearization in 3-dimensional topology*, Proc. I.C.M. (Helsinki, 1978), Acad. Sci. Fenn., Helsinki, 1980, pp. 463–468.

[Hat2] ______, *A proof of the Smale Conjecture*, Ann. of Math. **117** (1983), 553–607.

[Hen] D. Henderson, *Relations between spaces of simplexwise linear and smooth embeddings*, Preprint.

[Spi] M. Spivak, *A comprehensive introduction to differential geometry*, Vol. 1, Publish or Perish, Wilmington, Del., 1979.

DEPARTMENT OF MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NEW YORK 14853