THE SIMPLE CONNECTIVITY OF THE SUM OF TWO DISKS

R. H. BING

1. Introduction. The following question was called to the author's attention several years ago by Eldon Dyer.

Question. Is the sum of two disks simply connected if their intersection is connected?

Later, the author saw a communication in which an erroneous proof was given that Example 1 of the present paper is not simply connected. We show in § 2 that Example 1 is simply connected. However, we give some examples (Examples 2, 3, 4) in §§ 3, 4, 5 that are not simply connected.

A topological characterization is given in § 4 of intersections that will prevent closed curves which finitely oscillate between two disks from being shrunk. If the intersection is snake-like or arcwise connected, such finitely oscillating curves can always be shrunk but there are examples in which infinitely oscillating curves cannot. It is the topology of the intersection which prevents the sum of two disks from being simply connected rather than the embeddings of the intersection in the disks as shown in §§ 4 and 5. In fact, as pointed out in § 6, much of what we have learned about the sums of disks applies to the sums of continuous curves.

We use Example 1 in § 7 to construct a peculiar group and show that a certain relation kills it.

All sets treated in this paper are metric.

Let I^n denote an *n-cell* and *Bd* I^n its *boundary*. A set A is *n-connected* if each map (continuous transformation) f of *Bd* I^{n+1} into A can be extended to map I^{n+1} into A. We say that $f(Bd\ I^{n+1})$ *can be shrunk to a point* if the map can be extended. A set is called an *ε-set* if its diameter is less than or equal to ε. A set A is *n-ULC* if for each $\varepsilon > 0$ there is a $\delta > 0$ such that each map of *Bd* I^{n+1} onto a δ-subset of A can be shrunk to a point on an ε-subset. A compact continuum is called a *continuous curve* if it is 0-*ULC*. A set is *simply connected* if it is 1-connected. It is *uniformly locally simply connected* if it is 1-*ULC*. We shall not treat higher types of connectivity in this paper.

We find it convenient to consider an abstract disk D rather than the square I^2. A map of *Bd* D is a *closed curve*. If h is a homeomorphism, $h(BdD)$ is a *simple closed curve*.

Received March 6, 1963. Research supported by NSF grants NSF-G21514 and NSF-G11665.

We shall use cylindrical coordinates (ρ, θ, z) to describe examples in E^3 (Euclidean 3 space). If no z coordinate is given, it is understood that $z = 0$. When we use D alone without subscripts it is understood that we mean the unit disk ($\rho \leqq 1$) in the $z = 0$ plane.

Let f be a map of $Bd\ D$ into E^2 so that the ρ value of each point of $f(Bd\ D)$ is positive. Let $k(\theta)$ be a map of the reals into the reals such that $k(\theta)$ mod 2π is the θ value of $f(1, \theta)$. We say that f *circles the origin n times* if $k(2\pi) - k(0) = 2\pi n$.

2. A false example. Let a, b be fixed numbers with $0 < a < b < 1$ and K_1 be a spiral connecting the circles $\rho = a$ and $\rho = b$ as shown in Figure 1 and given by the following formula.

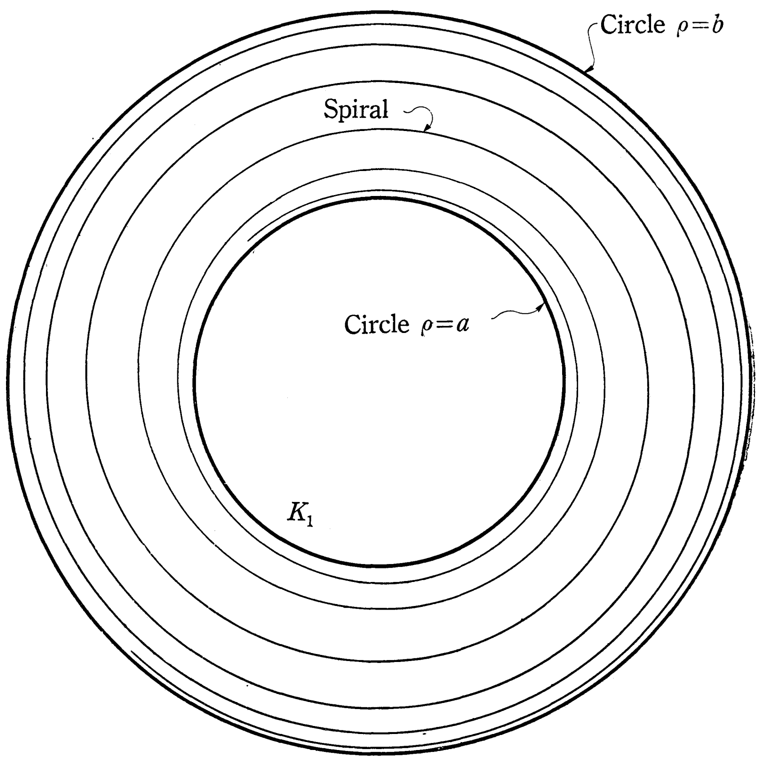


Figure 1

$$K_1 = \{(\rho, \theta)/\rho = a, b, \text{ or } (b + ae^{\theta})/(1 + e^{\theta})\} .$$

M. K. Fort showed [3] that any bounded plane continuum which has K_1 as a continuous image separates the plane.

EXAMPLE 1. Let D_1 be a disk in E^3 defined by

$$D_1 = \{(\rho, \theta, z)/\rho \leqq 1, z = \text{distance from } (\rho, \theta, 0) \text{ to } K_1\} .$$

Let D_2 be the reflection of D_1 through the $z = 0$ plane. Then $D_1 + D_2$ is the sum of two disks whose intersection is the connected set K_1.

THEOREM 1. *Example* 1 *is simply connected.*

Proof. Let f be a map of $Bd\ D$ into $D_1 + D_2$. We show that $D_1 + D_2$ is simply connected by showing that f can be extended to map D into $D_1 + D_2$.

With no loss of generality we suppose that the ρ value of each point of $f(Bd\ D) \geqq a$.

Special case. (The θ value of each point is fixed under f and the ρ value of each point of $f(Bd\ D) < b$.) In this special case we start by extending f to the circle $\rho = a$ by insisting that f is fixed on this circle.

For each point $q = (1, \theta_q)$ of $Bd\ D$ such that $f(q) \in K_1$, let S_q be the spiral from q about the circle $\rho = a$ described by the formulas $\rho \leqq 1$, $\rho = (2 - a + ae^{(\theta - \theta_q)})/(1 + e^{(\theta - \theta_q)})$, $\theta_q \leqq \theta$. Let f be extended to map S_q into K_1 so that f preserves the θ value of each point of S_q. This extension is made for each such spiral S_q for each point q of $Bd\ D$ such that $f(q) \in K_1$. Note that we have mapped a closed subset of D into K_1 and each component of $D - f^{-1}(K_1)$ other than the interior of $\rho = a$ intersects $Bd\ D$ in an open arc.

Let g_1 be the map of $f^{-1}(K_1 + D_1 \cdot f(Bd\ D)) = f^{-1}(D_1)$ into D_1 given by extended f. Then g_1 can be extended to take D into D_1. For convenience we also call this extended map g_1. Similarly there is a map g_2 of D into D_2 such that $g_2 = f$ on $f^{-1}(K_1 + D_2 \cdot f(Bd\ D)) = f^{-1}(D_2)$. Let g be the map of D into $D_1 + D_2$ given by g_1 on each component of $D - f^{-1}(K_1)$ which has an arc which goes into D_1 under f and $g = g_2$ on the rest of D.

Less special case. (f circles the origin once and the ρ value of each point of $f(Bd\ D)$ is less than b.) We show that there is a homotopy $h_t(0 \leqq t \leqq 1)$ of $Bd\ D$ into $D_1 + D_2$ such that $h_0 = f$, h_1 preserves the θ value of each point of $Bd\ D$ while the ρ value of each point of $h_1(Bd\ D)$ is less than b. The less special case then follows from the special case.

Let $k(\theta)$ be the function that shows that f circles the origin once. For convenience we suppose that $k(0) = 0$, $k(2\pi) = 2\pi$. Let $k_t(\theta) = t\theta + (1 - t)k(\theta)$, $(0 \leqq t \leqq 1)$. As t goes from 0 to 1, $k_t(\theta)$ goes from $k(\theta)$ to θ. For each point $p = (1, \theta_p)$ of $f^{-1}(K_1)$ we define $h_t(p)$ as a point in K_1 so that the θ value of $h_t(p)$ is $k_t(\theta_p)$. The ρ value of $h_t(p)$ is uniquely determined since the three arc components of K_1 are 1-manifolds almost normal to lines through the origin.

The homotopy h_t on $f^{-1}(K_1)$ is extended to $Bd\ D$ so that $h_t(p) \in D_i$ $(i=1, 2)$ if $f(p) \in D_i$, h_1 preserves the θ value of points of BdD and the value of each point of $h_1(Bd\ D)$ is less than b.

The following version of the less special case follows by a similar argument.

Alternative less special case. (f circles the origin once and the value of each point of $f(Bd\ D)$ is greater than a.)

General case. We suppose that $f(Bd\ D)$ intersects the spiral of K_1 in at least three points. Subdivide $Bd\ D$ into arcs $x_1x_2, x_2x_3, \cdots, x_nx_1 (n \geqq 3)$ so that no $f(x_ix_{i+1})$ (addition on subscripts is mod n so that $x_nx_{n+1} = x_nx_1$) intersects both circles in K_1 but each $f(x_i)$ is on the spiral of K_1. Let $x_iz_ix_{i+1}$ be the chord in D from x_i to x_{i+1}.

Extend the map f of $Bd\ D$ into $D_1 + D_2$ to map the chord $x_iz_ix_{i+1}$ into D_1 so that $f(x_iz_ix_i)$ misses the circles in K_1 and f on $x_ix_{i+1} + x_iz_ix_{i+1}$ circles the origin once. It follows from applications of the less special case and its alternative form that we can extend f to take the interiors of the $(x_ix_{i+1} + x_iz_ix_{i+1})$'s into $D_1 + D_2$. We can then extend f to the disk in D bounded by the chords into D_1.

3. **A true example.** Let C be a Cantor set on the numbers between 1/2 and 1. Let K_2 be the set in the plane consisting of the sum of circles in the plane with centers at the origin and radii in C and spirals joining adjacent circles as shown in Figure 2 and given by the following formula.

$$K_2 = \{(\rho, \theta)/\rho \in C \text{ or } \rho = (b + ae^{\theta})/(1 + e^{\theta})\}$$

where a, b are adjacent numbers of C with $a < b$.

EXAMPLE 2. Let E_1 be the disk in E^3 defined by

$$E_1 = \{(\rho, \theta, z)/\rho \leqq 1, z = \text{distance from } (\rho, \theta, 0) \text{ to } K_2\}\ .$$

Let E_2 be the reflection of E_1 through the plane $z = 0$. Then $E_1 + E_2$ is the sum of two disks whose sum is the connected set K_2.

Before proving that Example 2 is not simply connected we investigate an interesting property of K_2. M. K. Fort, Jr. showed [3] that any compact continuum in the plane separates the plane if it maps onto K_1. We modify his argument slightly to show the following.

THEOREM 2. *If f maps a closed bounded connected subset of the*

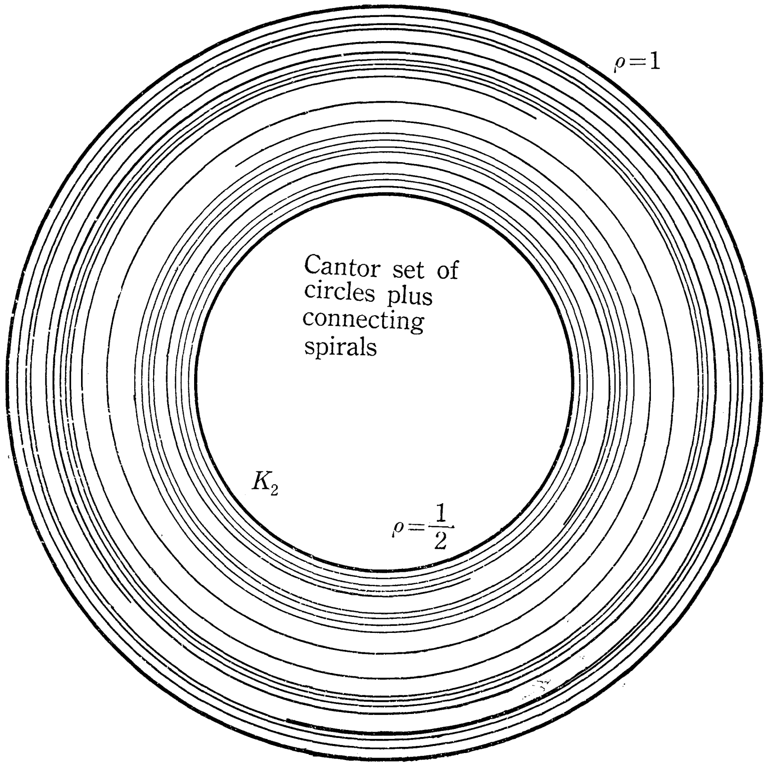


Figure 2

plane onto K_2 *then for each circle* J *in* K_2, *each component of* $f^{-1}(J)$ *separates the plane.*

Proof. Let S^1 be the circle $\rho = 1$ and define g: $K_2 \rightarrow S^1$ by $g(\rho, \theta) = (1, \theta)$. It is easy to verify that (K_2, S^1, g) is a locally trivial fiber space with totally disconnected fibers.

Suppose X is a component of $f^{-1}(J)$ that does not separate the plane. There is a homotopy pulling the map gf of X into S^1 to a constant map. Since S^1 is an ANR there is a neighborhood N of X such that the map gf of $\bar{N} \cdot f^{-1}(K_2)$ into S^1 is homotopic to a constant map. Take N so close to X that $\bar{N}$ does not cover $f^{-1}(K_2)$.

Let Y be a continuum in $f^{-1}(K_2)$ irreducible from $E^2 - N$ to X. Note that $Y \subset \bar{N}$, $Y \cdot X \neq 0$, and $Y \not\subset X$. It follows from the lemma on page 542 of [3] that $f(Y)$ is contained in an arc component of K_2. This violates the condition that the arc component of K_2 containing J does not intersect $K_2 - J$.

THEOREM 3. *Example* 2 *is not simply connected.*

Proof. Let x and y be points on the inner and outer circles in

K_2 and xz_iy be an arc in E_i from x to y. Let f be a map of $Bd\ D$ onto $xz_1y + xz_2y$ so that the upper half of $Bd\ D$ goes homeomorphically onto xz_1y and the lower half of $Bd\ D$ goes homeomorphically onto xz_2y. We show that Example 2 is not simply connected by showing that f cannot be extended to map D into $E_1 + E_2$.

Assume that f can be extended to send all of D into $E_1 + E_2$. We show that under this false assumption that $p = (1, 0, 0)$ and $q = (1, \pi, 0)$ belong to the same component of $f^{-1}(K_2)$. If they did not belong to the same component, it follows from Theorem 14 on page 171 of [6] (Theorem 10 on page 185 of 1932 edition) that there is a simple closed curve J in the plane $z = 0$ which misses $f^{-1}(K_2)$ and separates p from q in this plane. There would then be an arc A in $J \cdot D$ that intersects both the upper and lower halves of $Bd\ D$. This is impossible since f takes the upper half of $Bd\ D$ into E_1 and the lower half into E_2 but no point of A into $E_1 \cdot E_2 = K_2$.

Let Y be the component of $f^{-1}(K_2)$ containing p and q. Let Z be a subcontinuum of Y irreducible from p to q. Note that f maps Z onto K_2.

If F is a subcontinuum of Z which separates the plane E^2, no bounded component of $E^2 - F$ intersects Z since Z is irreducible from p to q and neither p nor q is in a bounded component of $E^2 - F$. Hence Z does not contain uncountably many mutually exclusive subcontinua each of which separates E^2. This contradicts Theorem 2 which says that for each circle J in $K_2, Z \cdot f^{-1}(J)$ separates E^2.

4. Finitely oscillating curves. A map of a simple closed curve J into the sum of two disks has only *finite oscillation* with respect to the two disks if J is the sum of a finite number of arcs such that the image of each lies in one of the disks. In some examples (Examples 3, 4, 5 to follow) finitely oscillating curves can be shrunk to points but some others cannot. The proof of Theorem 3 showed that Example 2 contained a finitely oscillating curve which could not be shrunk to a point.

We shall show that whether or not all finitely oscillating curves in the sum of two disks can be shrunk to points in the sum is dependent on whether or not the intersection has a certain extremal inverse property. A set X has the *extremal inverse property* with respect to its points p, q if there is a continuum Z in disk D with points p', q' on $Bd\ D$ and a map of Z into X that takes p', q' to p, q respectively.

Let K_3 be the sum of a triod T and a spiral S about T as shown in Figure 3 and given by the following equations.

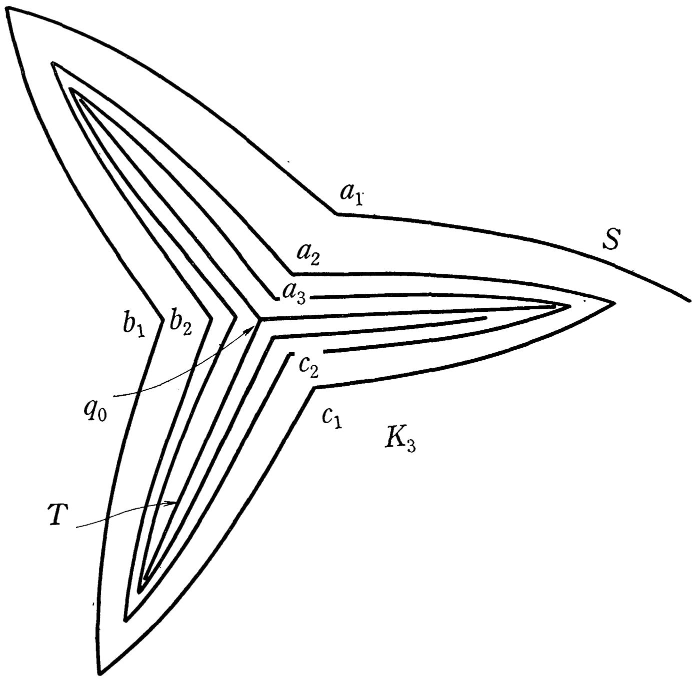


Figure 3

$$T = \{(\rho, \theta)/\rho \leqq 1, \theta = 0, 2\pi/3, \text{ or } 4\pi/3\}\,,$$
$$S = \{(\rho, \theta)/\rho = |\cos 3\theta/2|^{\theta} + 1/\theta, \theta \geqq 2\pi\}\,.$$

EXAMPLE 3. Let D_1, D_2 be two disks whose intersection is K_3.

THEOREM 4. *K_3 has the extremal inverse property with respect to each pair of its points.*

Proof. We consider only the case where $p \in T$ and $q \in S$. Let S' be another spiral about T which misses S, f be a retraction of $S' + T$ onto T, and p' be a point of S' that maps onto p under f. Extend f to the identity on $S + T$. There is a disk containing $T + S + S'$ which has p' and q on its boundary.

THEOREM 5. *Each snake-like continuum has the extremal inverse property with respect to each pair of its points.*

Proof. Apply the following result to a subcontinuum of the snake-like continuum irreducible between the two points under consideration.

THEOREM 6. *Each snake-like continuum is the image of a pseudo-arc.*

Proof. This theorem has been proved by each of Fearnley [2], Lelek [4], and Mioduszewski [5] but we include a slightly different proof.

Let $D_1, D_2, \cdots$ be a sequence such that D_i is a $1/i$ chain properly covering snake-like continuum X and such that D_{i+1} is a refinement of D_i. It follows from Theorem 7 of [1] that if P is a pseudo-arc there is a sequence of proper open coverings $E_1, E_2, \cdots$ of P such that E_i has the same number of links as D_i and for the jth link of D_{i+1} there is an integer $n(i, j)$ such that the jth link of D_{i+1} lies in the $n(i, j)$th link of D_i and the jth link of E_{i+1} lies in the $n(i, j)$th link of E_i.

For each point p of P let $e(p, i)$ be the sum of the links of E_i containing p and $d(p, i)$ be the sum of the corresponding links of D_i. Note that $e(p, i+1) \subset e(p, i)$ and $d(p, i+1) \subset d(p, i)$. For each point p of P let $f(p)$ be the intersection of the closures of $d(p, i)$'s. Then f is a continuous transformation of P onto X.

THEOREM 7. *If a set has the extremal inverse property, so does each of its continuous images.*

THEOREM 8. *Each arcwise connected set has the extremal inverse property with respect to each pair of its points.*

Note that the following theorem applies to simply connected and uniformly locally simply connected continuous curves as well as merely to disks.

THEOREM 9. *Let A_1, A_2 be two compact sets each of which is* 0-*connected,* 1-*connected,* 0-ULC, *and* 1-ULC. *A necessary and sufficient condition that each finitely oscillating curve with respect to A_1, A_2 can be shrunk to a point in $A_1 + A_2$ is that $A_1 \cdot A_2$ has the extremal inverse property with respect to each pair of its points.*

Proof. If $A_1 \cdot A_2$ does not have the extremal inverse property with respect to point x, y of $A_1 \cdot A_2$, let f be a map of $Bd\, D$ into $A_1 + A_2$ such that the upper half of $Bd\, D$ goes into a path in A_1 from x to y and the lower half of $Bd\, D$ goes into a path in A_2 from x to y. It follows from the proof of Theorem 3 that $f(Bd\, D)$ cannot be shrunk to a point in $A_1 + A_2$.

To prove the sufficiency case consider a map f of $Bd\, D$ into $A_1 + A_2$ so that $Bd\, D$ is the sum of arcs $x_1x_2, x_2x_3, \cdots, x_nx_1$ so that each $f(x_i)$ lies on $A_1 \cdot A_2$ and each $f(x_ix_{i+1})$ lies in one of A_1, A_2. Just as we used chords in the general case of the proof of Theorem 1, we consider continua $Z_1, Z_2, \cdots, Z_n$ in D so that Z_i contains x_i and x_{i+1} and f can be extended to take $Z_1 + Z_2 + \cdots + Z_n$ into $A_1 \cdot A_2$. Then as in the proof of the General Case of the proof of Theorem 1 we extend f to take the components of $D-(Z_1+Z_2+\cdots+Z_n)$ which intersect $Bd\,D$ into the appropriate one of A_1, A_2 and then extend f to take the rest of D into A_1. To extend f to take D into A_i for example, one would add a null sequence of arcs in D to $Z_1+Z_2+\cdots+Z_n$ to get a set Z_0 so that $D-Z_0$ is a null sequence of open disks, use the fact that A_i is 0-connected and 0-ULC to extend f to Z_0, and finally use the fact that A_i is 1-connected and 1-ULC to extend f to take D into A_i.

THEOREM 10. *Suppose D_1, D_2 are two disks whose intersection is a continuum X and axb is an arc in D_1 that intersects X only at a and b. If f is a map of $Bd\,D$ into D_1+D_2 such that f takes the upper half of $Bd\,D$ onto axb and the lower half into D_2, then $f(Bd\,D)$ can be shrunk to a point in D_1+D_2.*

Proof. Since X has the extremal inverse property with respect to a and b as shown by its embedding in D_1, there is a continuum X' in D intersecting the inverses under f of a and b such that f may be extended to $X'+Bd\,D$. Then f is extended to take the part of D in component of $D-X'$ that contains upper arc of $Bd\,D$ into D_1 and the rest of D into D_2.

Question. The question suggests itself as to which continua have the extremal inverse property with respect to each pair of their points. Example 2 does not have it. Example 1 and 3 do. So do Examples 4 and 5 to be given in the next section. Perhaps Example 2 is unnecessarily complicated as an example of a continuum without the extremal inverse property in that it separates the plane into infinitely many pieces. Perhaps there is an example that does not separate the plane. Does each three branched tree-like plane continuum have the extremal inverse property with respect to each pair of its points? (A compact continuum is a three branched tree-like continuum if it is not snake-like but for each positive number ε it has an ε-cover whose 1-nerve is a triod.)

5. Infinite oscillation. We use rectilinear coordinates to define the two sets shown in Figures 4, 5.

$$K_4=\{(x,y)/(x=0,\ -2\leqq y\leqq 2),\ (y=1+\sin 1/x,\ 0<x\leqq 1)\,,\ \text{or } (y=-1+\sin 1/x,\ -1\leqq x<0)\}\,.$$

$K_5=$ sum of points of K_3 on or to the left of the vertical line through $(1,0)$ plus the interval from $(1,0)$ to $(1,1)$.

Theorems 5 and 8 show that K_4 and K_5 have the extremal

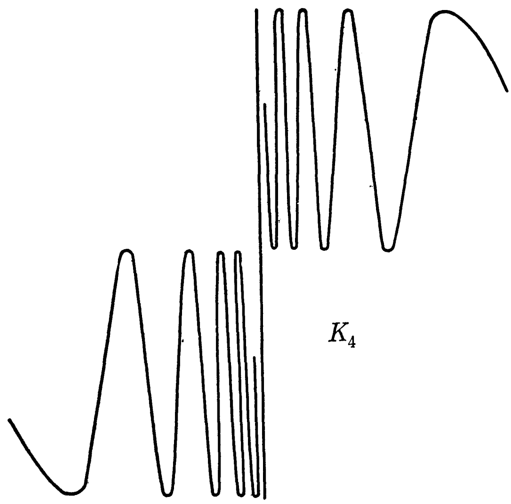


Figure 4

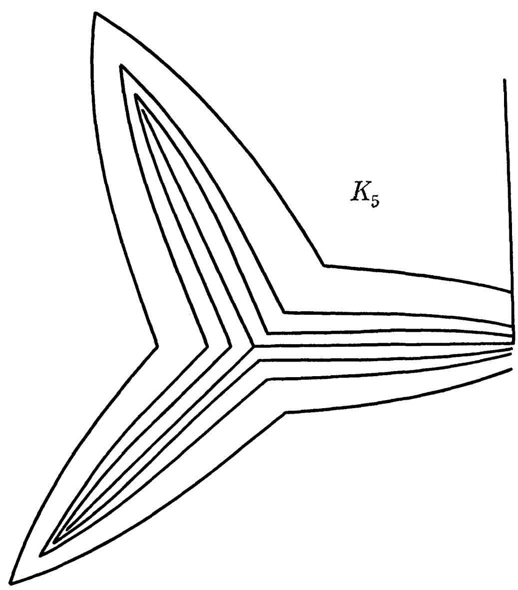


Figure 5

inverse property with respect to each pair of their points. It follows from Theorem 9 that finitely oscillating curves in the following two examples can be shrunk to points in the examples.

EXAMPLE 4. Two disks sewed together along K_4.

EXAMPLE 5. Two disks sewed together along K_5.

THEOREM 11. *Examples* 3, 4, 5 *are not simply connected.*

We only prove the first third of this theorem since the other parts are analogous. We suppose that the disks D_1, D_2 of Example 3 are obtained by pushing parts of a circular disk in the $z = 0$ plane up and down respectively as done in Examples 1 and 2. Theorem 13 shows that there is no loss of generality in supposing this. The disks would be larger than those in Examples 1 and 2 since K_3 is larger than those disks.

Proof that Example 3 *is not simply connected.* Let $a_1, a_2, \cdots$ be the points of K_3 on the open ray $\theta = \pi/3$ ordered inversely according to their distances from the origin q_0. Let $b_1, b_2, \cdots$ be the corresponding points of K_3 on the open ray $\theta = \pi$ and $c_1, c_2, \cdots$ be the corresponding points on the ray $\theta = 5\pi/3$.

Let p_i be the point of $Bd\, D$ whose θ value is $1/i$. Let p_0 be the point of $Bd\, D$ whose θ value is 0. Use p_ip_j to denote the arc on $Bd\, D$ in a clockwise direction from p_i to p_j.

Let f be a map of $Bd\, D$ into $D_1 + D_2$ satisfying the following conditions.

$$\begin{aligned}
&f(p_j) = q_0 \text{ (origin) } (j = 0, 2, 4, 6, \cdots)\,, \\
&f(p_{6i-5}) = a_i\,, \\
&f(p_{6i-3}) = b_i\,, \\
&f(p_{6i-1}) = c_i\,, \\
&f(p_{2i-2}p_{2i-1}) \subset D_1\,, \\
&f(p_{2i-1}p_{2i}) \subset D_2\,.
\end{aligned}$$

The θ value on each $f(p_ip_{i+1})$ is a constant and f takes the θ values of p_ip_{i+1} linearly onto the ρ values of $f(p_ip_{i+1})$.

Assume f can be extended to take D into $D_1 + D_2$. We call the extended map f. In this extension we suppose that no component of $f^{-1}(q_0)$ separates the $z = 0$ plane. (If a component X did separate the plane, we could modify f to map X plus each of its bounded complementary domains in $z = 0$ into q_0.)

Let F be the part of $D_1 + D_2$ whose ρ value is less than or equal to 1. A finite number of spanning arcs in D separates p_0 from $f^{-1}(D_1 + D_2 - F)$ in D so that no one of the arcs intersects $f^{-1}(q_0)$. Hence, we can cut down the disk D to a disk E such that E agrees with D in a neighborhood of $p_0, f(E) \subset F$, each point of $Bd\, E \cdot f^{-1}(q_0)$ lies on $Bd\, D$, and E agrees with D in a neighborhood of each such point. Let Z be the closure of $Bd\, E - Bd\, D$. Note that $q_0 \notin f(Z)$. We shall obtain a contradiction to the assumption that f can be extended to take D into $D_1 + D_2$ by showing that the map f of $Bd\, E$ into F cannot be extended to take E into F.

With no loss of generality we suppose that $p_0, p_1, p_2, \cdots$ all belong to $Bd\, E$. Note that no component of $E \cdot f^{-1}(K_3)$ intersects two p_i's unless perhaps they both have even subscripts since no two of $a_1, a_2, \cdots, b_1, b_2, \cdots, c_1, c_2, \cdots$ belong to the same component of $K_3 \cdot F$. For i odd, the component Y_i of $E \cdot f^{-1}(K_3)$ containing p_i separates $Bd\, E$ since $f(Bd\, E)$ crosses from D_1 to D_2 at $f(p_i)$. Since $Bd\, D$ intersects Y_i in at most a finite number of points and $f(Bd\, D)$ does not cross from D_1 to D_2 at the image of any of these points other than p_i, Y_i must intersect Z. Let q_i be a point of $Y_i \cdot Z$.

Note that since $a_1, a_2, \cdots, b_1, b_2, \cdots, c_1, c_2, \cdots$ belong to different components of $F \cdot K_3$, $Y_i \neq Y_j$ if i, j are different odd positive integers. Let q_∞ be a limit point of $q_1, q_3, q_5, \cdots$. Since for i sufficiently large Y_{i+2} separates Y_i from p_0 in E, $q_1, q_3, \cdots$ converges to q_∞ from the clockwise side. Since q_∞ is a limit point of each of ΣY_{6i-5}, ΣY_{6i-3}, ΣY_{6i-1}, then $f(q_\infty)$ is a limit point of each of $f(\Sigma Y_{6i-5})$, $f(\Sigma Y_{6i-3})$, $f(\Sigma Y_{6i-1})$. The only point common to the closures of these sets is the point q_0, so $f(q_\infty) = q_0$. However, $f(q_\infty) \neq q_0$ since $q_\infty \in Z$ and $q_0 \notin f(Z)$.

THEOREM 12. *The sum of two disks is simply connected if their intersection is connected and locally arcwise connected.*

Proof. Let f be a map of $Bd\, D$ into the sum of two disks D_1, D_2 such that $D_1 \cdot D_2$ is a continuous curve. For each arc ab of $Bd\, D$ which intersects $f^{-1}(D_1 \cdot D_2)$ only in its end points, extend f to map the chord acb of D into an arc in $D_1 \cdot D_2$ such that the diameter of $f(acb)$ is no more than twice the diameter of any other arc in $D_1 \cdot D_2$ from $f(a)$ to $f(b)$. Let f_i be a mapping of D into D_i that agrees with f on the part of D going into D_i under f. Then the extended f is f_1 on the components of D minus the chords which contain a point of $Bd\, D$ that f sends into $D_1 - D_2$ and is f_2 on the rest of D.

THEOREM 13. *The topology of the intersection of two disks determines whether or not their sum is simply connected.*

Proof. Suppose D_1, D_2, E_1, E_2 are disks and h is a homeomorphism of $D_1 \cdot D_2$ onto $E_1 \cdot E_2$. Let D be a circular disk and f a map of $Bd\, D$ into $D_1 + D_2$. We assume that $E_1 + E_2$ is simply connected and show that this assumption implies that f can be extended to map D into $D_1 + D_2$. We assume there are at least three points of $Bd\, D$ that f sends into $D_1 \cdot D_2$

Let g be a map of $f^{-1}(D_1 \cdot D_2 \cdot f(Bd\, D))$ into $E_1 \cdot E_2$ given by $g = hf$. For each arc ab of $Bd\, D$ which intersects $f^{-1}(D_1 \cdot D_2 \cdot f(Bd\, D))$ only in its end points, extend g to map the chord acb of D onto an arc in E_i if $f(ab) \subset D_i$ with the restriction that the diameter of

$g(acb)$ is not more than twice the diameter of any other arc in E_i from $g(a)$ to $g(b)$. Let E be the subdisk of D such that g has been defined to map $Bd\, E$ into $E_1 + E_2$.

Since $E_1 + E_2$ is simply connected, we extend g to map E into $E_1 + E_2$. Call the extension g. Consider $g^{-1}(E_1 \cdot E_2 \cdot g(E)) = X$. No two points of $Bd\, D$ can be joined by an arc in $D - X$ unless the points go into the same one of D_1, D_2 under f.

Define f on X to be $h^{-1}g$. Let f_i be the extended f restricted to $f^{-1}(D_i \cdot f(X + Bd\, D))$. Extend f_i to map D into D_i and call the extended map f_i. The extended map f is f_1 on each component of $D - X$ which has points of $Bd\, D$ which are sent by f into D_1 and is f_2 on the rest of D.

6. **Adding continuous curves.** What we have learned about the sum of disks partially applies to the sums of other continua. If the intersection of two disks is so bad as to make the sum of the disks not simply connected, it is bad enough to keep any two continuous curves whatever with the same intersection from being simply connected. The following example illustrated in Figure 6 shows that the converse is not true.

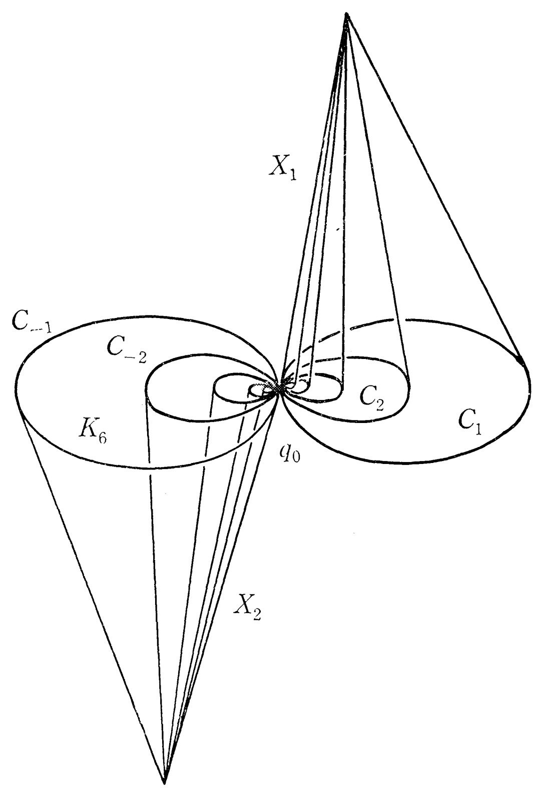


Figure 6

EXAMPLE 6. Let $C_i\,(i = 1, 2, 3, \cdots)$ be the circle in the x, y plane with equation $(x - 1/i)^2 + y^2 = (1/i)^2$. Denote the origin by q_0. Let $K_6 = C_1 + C_2 + \cdots + q_0 + C_{-1} + C_{-2} + \cdots$. The cone $X_1 \cdot$ over $C_1 + C_2 + \cdots + q_0$ from a point above the xy plane is simply connected as is the cone X_2 over $q_0 + C_{-1} + C_{-2} + \cdots$ from a point below the xy plane. Although $X_1 \cdot X_2$ is a point, $X_1 + X_2$ is not simply connected.

THEOREM 14. *Suppose D_1, D_2 are two disks and F_1, F_2 are two continuous curves such that $D_1 \cdot D_2$ is homeomorphic with $F_1 \cdot F_2$. Then $D_1 + D_2$ is simply connected if $F_1 + F_2$ is.*

Proof. The proof is the same as the proof of Theorem 13 except that g maps $Bd\,E$ into $F_1 + F_2$ instead of into $E_1 + E_2$.

THEOREM 15. *Suppose G_1, G_2, G_3, G_4 are four simply connected and uniformly locally simply connected continuous such that $G_1 \cdot G_2$ is topologically equivalent to $G_3 \cdot G_4$. Then the fundamental group of $G_1 + G_2$ is isomorphic to the fundamental group of $G_3 + G_4$.*

Proof. Whether or not a loop in $G_1 + G_2$ can be shrunk to a point depends on how it crosses back and forth between G_1 and G_2. Suppose h is a homeomorphism of $G_1 \cdot G_2$ onto $G_3 \cdot G_4$ and x_0 is a point of $G_1 \cdot G_2$ that acts as a starting point of loops in $G_1 + G_2$ to determine the fundamental group of $G_1 + G_2$. We use $h(x_0)$ as a starting point for the loops in $G_3 + G_4$ to determine the fundamental group of $G_3 + G_4$.

Let $\{f\}$ be an element of the fundamental group of $G_1 + G_2$. It is an equivalence class of maps of the interval $[0, 1]$ into $G_1 + G_2$ such that the ends of $[0, 1]$ go into x_0. Let f be an element of $\{f\}$. Let f' be a map of $[0, 1]$ into $G_3 + G_4$ such that

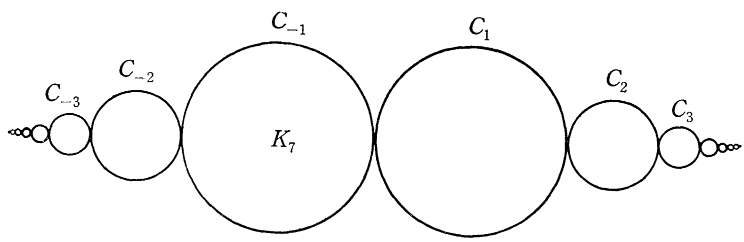
$$f'(x) = hf(x) \text{ if } f(x) \in G_1 \cdot G_2\,,$$
$$f'(x) \in G_{i+2} \text{ if } f(x) \in G_i\,.$$

Although these two conditions do not precisely define f', any two maps satisfying this condition are homotopic in $G_3 + G_4$. The element of the fundamental group of $G_3 + G_4$ corresponding to the element $\{f\}$ of the fundamental group of $G_1 + G_2$ is the equivalence class of loops containing f'.

Question. The preceding theorem suggests a topological invariant of compact closed sets. Two sets A, B are alike in a certain sense provided the sum of two Hilbert cubes sewed together along A have the same fundamental group as the sum of two Hilbert cubes sewed together along B. Is there a simpler characterization of this property?

7. An interesting group. One might attempt to compute the fundamental group of Example 1 by cutting it into two pieces with a vertical plane through the origin, fatten each piece to make them intersect in an open subset of their sum, find the fundamental group of each piece, and then apply Van Kampen's theorem to get the fundamental group of Example 1. We ignore the fattening since, being equivalent to taking the slice slightly to one side of the origin, it does not change the fundamental group of the pieces.

Each piece can be folded like a fan and deformed onto a set topologically equivalent to a set K_7 shown in Figure 7 and defined as follows.



Fignre 7

$$K_7 = \text{closure of } (C_1 + C_2 + \cdots + C_{-1} + C_{-2} + \cdots)$$

where C_i is the circle in the xy plane with $[(i-1)/i,\ i/(i+1)]$ as diameter and C_{-i} is the circle with $[(-(i-1)/i,\ -i/(i+1)]$ as diameter. (We used $[a, b]$ to denote the interval on the x axis from a to b.) The fundamental group of each piece into which we cut Example 1 is the same as the fundamental group of K_7.

Consider the origin as the starting point of loops in K_7 to determine its fundamental group $G(K_7)$. Then a loop is a map of the interval $[0, 1]$ into K_7 that sends the ends of the interval to the origin and an element of $G(K_7)$ is an equivalence class of loops. We can associate words with loops. If a loop goes across the top semicircle of C_i from left to right we write i; if it goes across this semicircle from right to left we write $\bar{i}$ and say i inverse. We call i and $\bar{i}$ letters and say that the letter is positive or negative according as i is positive or negative. Since we are permitting C_i's with negative subscripts, i inverse differs from $-i$. The inverse of $\bar{i}$ is i. A loop then corresponds to an ordered collection of letters (called a word) with the following restrictions.

a. No letter appears in any word more than a finite number of times.

b. There is not infinite oscillation between positive and negative letters.

Let us consider what words are equated if two pieces into which we divided Example 1 are joined together again. If a loop is slid from one piece to the other until it comes back to the first in one direction, each letter i (or $\bar{i}$) in it has been changed to $i+1$ (or $\overline{i+1}$) and if the loop is slid in the other direction, these are replaced by $i-1$ (or $\overline{i-1}$). Since we skipped 0 in putting subscripts on the C_i's we suppose $-1+1=1$ and $1-1=-1$. When we replace each i or $\bar{i}$ in a word W by $i+1$ or $\overline{i+1}$, we have produced a right shift and call the new word $R(W)$. We note that if $W_1 = R(W_2)$, then W_2 may be obtained from W_1 by a left shift and say $W_2 = \bar{R}(W_1)$.

Let us change the group G {equivalence classes of W_α's} by also putting words in the same equivalence class if they are equivalent after a shift. This shifting operation is to be permitted in equating words only a finite number of times as opposed to cancellation which was permitted infinitely often. We call the resulting group G {equivalence classes of W_α's$/R(W_\alpha) = W_\alpha$}. Since the fundamental group of Example 1 is trivial, it follows that this group is trivial.

The inverse of a word is obtained by reversing the order of the letters and replacing each letter with its inverse. If in a word there appears two adjacent subwords which are inverses of each other, the word obtained by canceling the subwords belongs to the same equivalence class with the original word. Infinite cancellation is permitted so that for example $(1, \bar{1}, 2, \bar{2}, 3, \bar{3}, \cdots)$ is equivalent to the trivial word.

Two words are equivalent if and only if they can be cancelled down to a common word. (We could have given more extensive rules but they boil down to this.) To multiply two words, we write one after the other. If $\{W_\alpha\}$ denotes the collection of words and G {equivalence classes of W_α's} denotes the group of equivalence classes of words, then

$$G(K_7) = G\,\{\text{equivalence classes of } W_\alpha\text{'s}\}\,.$$

To show algebraically that G {equivalence classes of W_α's$/R(W_\alpha) = W_\alpha$} is trivial, consider a word W. Since we did not permit infinite oscillation between positive and negative letters of W, we can express W as $W_1 W_2 \cdots W_n$ where each W_i has either all positive or all negative letters. We show that W is trivial by showing that each W_i is. We consider only the case where W_i consists of positive letters since the other case is analogous.

Consider $X = \bar{W}_i R(\bar{W}_i) R^2(\bar{W}_i) R^3(\bar{W}_i) \cdots$. It is a word since it only contains positive letters and none appears more than a finite number of times. Then

$$W_i = W_i X \bar{X} = R(\bar{W}_i) R^2(\bar{W}_i) \cdots \bar{X} = R(X)\bar{X} = X\bar{X} = 1 .$$

One might wonder what would have happened if we had not imposed the condition that there is not infinite oscillation between the positive and negative letters in words. This would have been equivalent to the fundamental group of K_7 after the sum of the bottom simicircles were shrunk to a point. Even after a shift, it seems that the group is not killed. After the shift we would have the fundamental group of Example 1 if the annulus in D_2 outside $\rho = a$ is shrunk to the circle $\rho = a$.

REFERENCES

1. R. H. Bing and F. B. Jones, *Another homogeneous plane continuum*, Trans. Amer. Math. Soc., **90** (1959), 171-192.
2. L. Fearnley, *Continuous mappings of the pseudo-arc*, Notices of the American Mathematical Society Abstract 564-171, **6** (1959), 841.
3. M. K. Fort, Jr., *Images of plane continua*, Amer. J. Math., **81** (1959), 541-546.
4. A. Lelek, *On weakly chainable continua*, Fundamenta Mathematicae, **51** (1962), 271-282.
5. J. Mioduszewski, *A functional conception of snake-like continua*, Fundamenta Mathematicae, **51** (1962), 179-189.
6. R. L. Moore, *Foundations of Point Set Theory*, Amer. Math. Soc. Colloq. Publ., vol. 13, revised edition, 1962.

THE UNIVERSITY OF WISCONSIN
THE INSTITUTE FOR ADVANCED STUDY