A THREE COLOR PROBLEM

by Gordon Simons[1] and Yi–Ching Yao

University of North Carolina and Colorado State University

ABSTRACT

It is well-known that the problem of determining whether a planar map is three-colorable is NP-complete. There are also several isolated results on three-colorability for certain classes of maps. Herein, we consider a new class of maps that consists of those involving lines and half-lines. It is shown that such maps are three-colorable if they satisfy a simple condition. An extension to curved lines is also considered.

[1]This author was partially supported by the National Science Foundation, Grant No. DMS–8902973.

1. Introduction. The problem of coloring a planar map has a long history. The famous four color conjecture, first posed by Francis Guthrie circa 1852, remained open for over a century and finally was proved by Appel and Haken (1976) with the aid of a computer. Besides the four color theorem, it is also well-known that a map is two-colorable if and only if every vertex of the map is of even degree. See, for instance, Theorem 2.3 of Saaty and Kainen (1977).

A simple characterization of three-colorability for planar maps has not been found. A plausible explanation for this failure was provided by Stockmeyer (1973) when he established "NP-completeness" for the problem of determining when a planar map is three-colorable. Thus the difficulty of the task of deciding three-colorability is on a par with that of finding solutions to the celebrated "travelling salesman problem", a problem for which no polynomial-time solution is known (or anticipated). Garey, Johnson and Stockmeyer (1976) show that the problem remains NP-complete even when attention is restricted to maps with countries having at most four sides.

One can still hope to establish three-colorability for maps with special properties. An interesting known result of this type concerns cubic maps, which are maps with vertices solely of degree three. Theorem 2.5 of Saaty and Kainen shows that a cubic map is three-colorable if and only if each country (region) is bounded by an even number of sides (boundaries).

Grötzsch (1958) proves the interesting result that a map is three-colorable if its dual (planar) graph contains no cycle of length three. (Informally, the dual graph of a map is formed by choosing a point in the interior of each country and joining any two such points with an edge whenever the corresponding countries have a common boundary.) Grünbaum (1963) generalizes this result to show that a map is three-colorable if its dual graph contains at most three cycles of length three. Counterexamples arise if the number of such cycles is allowed to exceed three.

There is an elegant result due to Brooks (1941) (see page 137 of Saaty and Kainen (1977)) which, when specialized to three colors, says that practically all "triangular maps" are three-colorable (see Theorem 2.6 of Saaty and Kainen). A map is *triangular*, according to Saaty and Kainen, "if every region has three sides." (Whether this applies to *unbounded* regions is not made clear. Brook's theorem is concerned with coloring the nodes of graphs; it says something quite precise about coloring the nodes of the dual graph of a map. The interested reader can tease out from this a precise translation to the context of maps.)

Readers interested in the general problem of three-colorability may wish to read Chapter 13 of Ore (1967).

A rather simple class of maps consists of those involving (straight, unbounded) lines in the Euclidean plane. Such maps are easily seen to be two-colorable. A natural extension includes maps involving lines and half-lines. This paper is an attempt to determine when such maps are three-colorable. In the next section, we first present a map involving three half-lines which is not three-colorable. We then describe a mild condition on the half-lines which guarantees three-colorability. In Section 3, we indicate how the results can be extended to maps involving unbounded curves.

When a half-line has its base point embedded in another (half-)line, a cycle of length three is formed in the dual graph. The number of these cycles roughly equals the number of half-lines. Therefore, our results can be viewed as complementing the work of Grötzsch and Grünbaum, which severely limits the allowable number of length-three cycles.

2. Main results. The map in Figure 1a below involves three half-lines which yield four regions in the Euclidean plane that are adjacent to one another. So it can not be three-colored. A more subtle example needing four colors is shown in Figure 1b. These examples show that a map made up of lines and half-lines is in general not three-colorable.

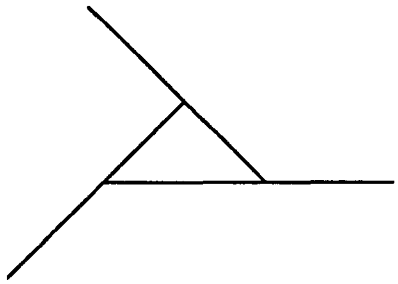

Figure 1a

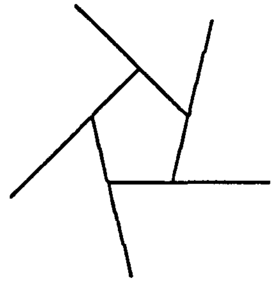

Figure 1b

The finite endpoint of a half-line will be referred to as its *base point.*

There are several useful restrictions to impose on the half-lines which can be made without affecting the generality of our results:

(a) If a half-line does not intersect any other (half-)line, remove it; it can not affect the coloring of the map.

(b) If the base point of a half-line is not embedded in another (half-)line, then remove the "tail" extending beyond the nearest (half-)line; it can not affect the coloring of the map. For example, the map in Figure 2a becomes the map in Figure 2b.

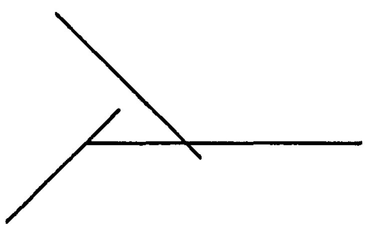

Figure 2a

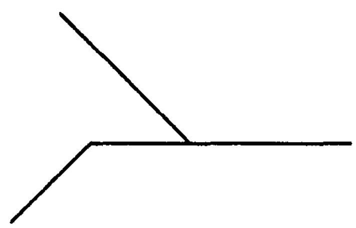

Figure 2b

(c) If a half-line is contained in another (half-)line, remove the former.

(d) If the union of two half-lines is a line, replace them by a single line. Thus it will be assumed that two (half-)lines intersect at no more than a single point.

We are now in a position to state our main condition and first theorem:

Condition A. *For each half-line* L, *there exists a line* L_0 *and half-lines* $L_1, \cdots, L_k$, *for some* $k \geq 1$, *such that* $L = L_k$, *and the base point of* L_i *is embedded in* L_{i-1} *for* $i = 1, \cdots, k$.

Theorem 1. *If a map satisfies condition A, then it is three-colorable.*

A discussion of Condition A and related topics must precede a proof of this theorem. (The stronger Theorem 2 below will eventually be proved.)

When a map satisfies Condition A, we can construct rooted trees to describe the relationships among the (half-)lines in the following sense. Each line is the root of exactly one tree, and each half-line is a node (non-root) of exactly one tree. The analogy of a family tree will be followed in that we shall refer to a (half-)line as the "father", and another half-line as the "son" if the base point of the latter is embedded in the former. (When three or more (half-)lines intersect at a common point, the construction of trees, and the paternity assignments, might not be unique.)

Some simple examples of rooted trees appear in Figures 3a, 4, 5a and 5b below. (The numbering and cross-hatching of some of the regions should be ignored for now.)

Arrows are shown on all (half-)lines of these figures. A half-line has a natural direction, viewed as *directed away from the base point*; a line does not. The directions of lines can be assigned arbitrarily.

Multiple sons are viewed as ordered, *according to their base points*, in the direction of the father's arrow. In Figures 3a and 4, L_1 is the "older", L_2 the "younger".

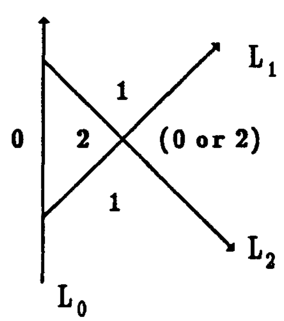

Figure 3a

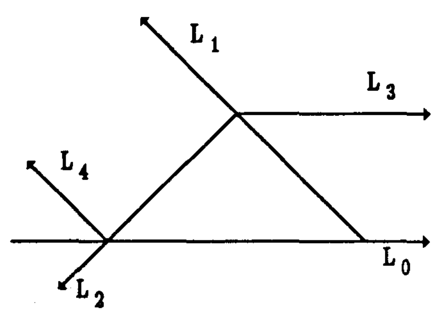

Figure 3b

Ambiguities can arise. Two are illustrated in Figure 3b. The half-lines L_2 and L_3 have a common base point in L_1. Either ("twin") may be viewed as L_1's older son. More confusing, L_4 can either be L_0's oldest son or L_2's only son (L_0's great-grandson). Either choice is satisfactory. Ambiguities can be finessed by perturbing the offending base points. But this is unnecessary.

We note that a map with only (whole) lines present can be two-colored since there is a simple way to update the colors of the regions when a new line is added: by reversing the colors of the regions on one side of the new line, and leaving the other side alone. But when half-lines are present, there is no obvious way to update the colors when a new half-line is added, assuming that it remains three-colorable. However, if the half-lines enter according to a proper order (in the sense of Condition A), then the updating can be accomplished by a simple updating procedure which tells which regions need to be updated and how. The proper order (a partial ordering) is for each father to enter before his sons, and his sons to enter from oldest to youngest.

When a half-line L enters, updating occurs within a wedge W(L), and the nature of the updating is determined by the product of two attributes $\alpha(L)$ and $\beta(L)$. The description of these concepts follows.

The wedge W(L), when L is a half-line, is the region bounded on one side by L and on another side by L's father, that portion of the L's father beyond the base point of L and *in the direction of its arrow.* By convention, when L is a line, W(L) will be defined as the region on the *right side of* L as one faces the arrow. The wedges for two half lines L_1 and L_2 are shown in Figure 4 below (perpendicular to their father for illustrative purposes only).

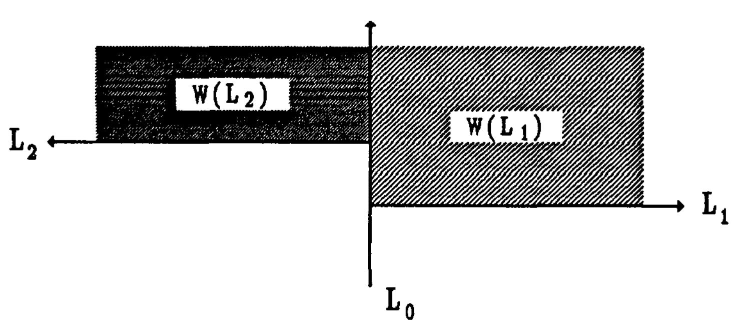

Figure 4

The *orientation* $\alpha(L)$ of a half-line L is 1 (–1) if, as one faces in the direction of the arrow of L's father, L is to the right (left). By convention, the value $\alpha(L) = -1$ will always be used when L is a line.

The parity of a (half–)line is defined recursively: By convention, the parity $\beta(L)$ of a line L is set equal to –1. If the parity of a (half–)line is $x \in \{-1,1\}$, then the parity of its oldest son is set equal to $-x$, the parity of the next oldest son is $+x$, etc.

The reader may verify for Figure 3a that $\alpha(L_0) = \beta(L_0) = -1$; $\alpha(L_1) = \beta(L_1) = 1$; $\alpha(L_2) = 1$ and $\beta(L_2) = -1$. For Figure 4, $\alpha(L_2) = \beta(L_2) = -1$.

It is possible to describe, *explicitly*, how to three-color the maps of Theorem 1: Let $\mathscr{L}$ be a family of (half-)lines which satisfies Condition A and generates a map denoted $M(\mathscr{L})$. A function $C : M(\mathscr{L}) \to \{0,1,2\}$ will be said to satisfy the *coloring property* if $C(R_1) \neq C(R_0)$ for every pair of regions R_0 and R_1 in $M(\mathscr{L})$ which shares a common boundary. We shall show that the function C defined by

$$C(R) = \left\{ \sum_{L \in \mathscr{L}} \alpha(L)\beta(L)\, I_{W(L)}(R) \right\} \pmod 3, \qquad R \in M(\mathscr{L}), \quad (1)$$

where

$$I_{W(L)}(R) = \begin{cases} 1 \text{ if } R \text{ is a subset of } W(L) \\ 0 \qquad\qquad \text{otherwise} \end{cases},$$

satisfies the coloring property.

The reader can easily verify that formula (1) produces the colors (numbers) shown in Figure 3a (the first number where two possibilities are shown), 5a and 5b.

When it comes to proving Theorem 1, it is possible to treat the various rooted trees independently: Each tree gives rise to a submap. If each submap can be three-colored, with colors identified by the numbers 0, 1 and 2, then the original map can be three-colored by adding mod 3 the colors associated with each region. A simple example of this process is shown below in Figures 5a,b,c. Figures 5a and 5b are three-colored submaps for the actual map shown in 5c; the numbers in 5c are found for each of its regions by adding mod

3 the corresponding numbers in 5a and 5b. The numbers shown in all three figures are those that would be produced by applying formula (1).

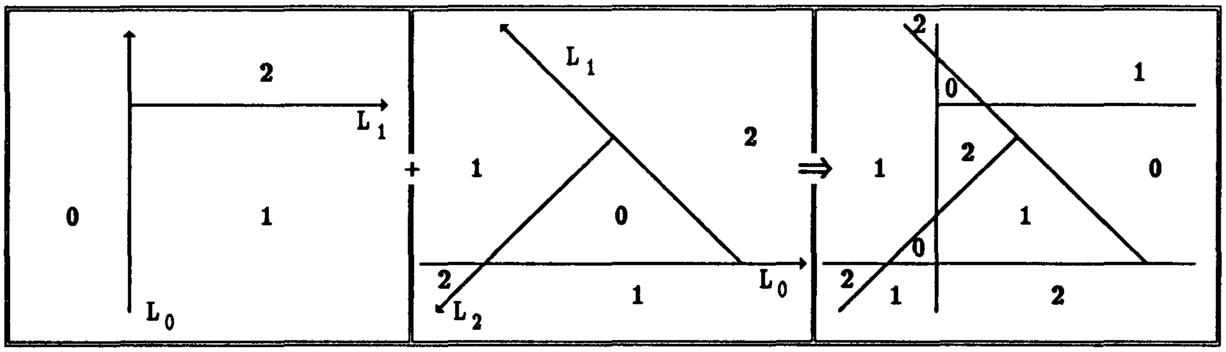

Figure 5a Figure 5b Figure 5c

Thus, we can focus attention on maps generated by a single rooted tree.

The function C defined in (1) satisfies something stronger than the coloring property. A function C: $M(\mathscr{L}) \to \{0,1,2\}$ will be said to satisfy the *parity property* if

$$C(R_1) - C(R_0) = \beta(L_1) \pmod 3 \quad (\neq 0), \tag{2}$$

whenever there exist a (half-)line L_0 and adjacent regions R_0 and R_1 on the right and left sides of L_0, respectively, relative to the direction of L_0's arrow, where L_1 is the youngest son of L_0 for which one of R_0 and R_1 is in the wedge $W(L_1)$, or it is L_0 if no such son exists.

The map shown in Figure 5a has the parity property: Let R_0, R_1, R_2 be the regions numbered 0,1,2, respectively. Then $C(R_0) - C(R_2) = -2 = \beta(L_1) \pmod 3$ with $\beta(L_1) = 1$, and $C(R_0) - C(R_1) = -1 = \beta(L_0) \pmod 3$ with $\beta(L_0) = -1$. Likewise, Figure 5b exhibits the parity property.

It is easily checked that *the parity property holds for $\mathscr{L}$ if it holds for each rooted tree of $\mathscr{L}$.*

The following implies Theorem 1 directly.

Theorem 2. *If a map* $M(\mathscr{L})$ *satisfies Condition A, then the function* $C: M(\mathscr{L}) \to \{0,1,2\}$ *described in* (1) *satisfies the parity property.*

Proof. It is enough to show this when $\mathscr{L}$ is a single rooted tree. The proof proceeds by induction based upon the number of (half-)lines in $\mathscr{L}$.

When $\mathscr{L}$ has a single line L_0, the parity condition is easily checked: The function C assigns the values 0 and 1 on the left and right sides of L_0, respectively, relative to the direction of its arrow, and, by conventions, $W(L_0)$ is the right side of L_0, and $\beta(L_0) = -1$. Thus $C(R_1) - C(R_0) = 0 - 1 = \beta(L_0)$, where $R_0 = W(L_0)$ and R_1 is its complement.

For the induction step, express $\mathscr{L}$ in the form $\mathscr{L} = \mathscr{L}_* + \{L_1\}$ (a disjoint union), where the half-line L_1 is both sonless and also the youngest son of a (half-)line $L_0 \in \mathscr{L}_*$. (As previously noted, $\mathscr{L}$ is partially ordered on the basis of paternity; L_1 is simply an extreme member of this partial ordering.) Clearly, the "reduced map" $M(\mathscr{L}_*)$ satisfies condition A. So by the induction assumption, the function C_* defined by

$$C_*(R) = \left\{ \sum_{L \in \mathscr{L}_*} \alpha(L)\beta(L)\, I_{W(L)}(R) \right\} (\text{mod } 3), \qquad R \in M(\mathscr{L}_*), \quad (3)$$

satisfies the parity property. Note that because of the way L_1 is chosen, the orientation and parity functions α and β in (3) do not require asterisks. Since C_* assigns the same color $C_*(R)$ to every point of a region $R \in M(\mathscr{L}_*)$, one may use the formula in (3) to describe the colors assigned by C_* to $R \in M(\mathscr{L})$. ($M(\mathscr{L})$ partitions the plane more finely than does the reduced map $M(\mathscr{L}_*)$.) Of course, it does not follow that C_* satisfies the parity property for the full map $M(\mathscr{L})$; it does not. But we have from (1) the relationship

$$C(R) = \{C_*(R) + \alpha(L_1)\beta(L_1)\, I_{W(L_1)}(R)\} (\text{mod } 3), \quad R \in M(\mathscr{L}). \qquad (4)$$

Let R_0 and R_1 be fixed adjacent regions of $M(\mathscr{L})$ on the right and left sides, respectively, of a (half-)line $L \in \mathscr{L}$ relative to the direction of L's arrow. Then,

$$C(R_1) - C(R_0) = \{C_*(R_1) - C_*(R_0) + \gamma(L_1)\} \pmod 3, \tag{5}$$

where

$$\gamma(L_1) := \alpha(L_1)\beta(L_1)\{I_{W(L_1)}(R_1) - I_{W(L_1)}(R_0)\}.$$

If $L = L_1$, then R_0 and R_1 are subsets of the *same* set R in the reduced map $M(\mathscr{L}_*)$ (so $C_*(R_1) = C_*(R_0)$), and exactly one of R_0 and R_1 is contained in the wedge $W(L_1)$. Hence, $\gamma(L_1) = \alpha(L_1)^2\beta(L_1) = \beta(L_1)$. So by (5), $C(R_1) - C(R_0) = \beta(L_1) \pmod 3$, which is the parity condition shown in (2) for this case.

If $L = L_0$, let L_2 be the youngest son of L_0 *within* $\mathscr{L}_*$ (thus excluding L_1) for which one of R_0 and R_1 is in the wedge $W(L_2)$; let $L_2 = L_0$ if there is no such son. Since the parity property holds for the reduced map $M(\mathscr{L}_*)$, $C_*(R_1) - C_*(R_0) = \beta(L_2) \pmod 3$. Thus

$$C(R_1) - C(R_0) = \{\beta(L_2) + \gamma(L_1)\} \pmod 3. \tag{6}$$

There are two cases to consider: (i) If R_0 and R_1 are both outside the wedge $W(L_1)$, then $\gamma(L_1) = 0$ in (6), which then becomes the appropriate form of the parity condition shown in (2) for this setting. (ii) If one of R_0 and R_1 is in the wedge $W(L_1)$, then $\beta(L_2) = -\beta(L_1)$ and $\gamma(L_1) = -\alpha(L_1)^2\beta(L_1) = -\beta(L_1)$. Then (6) says $C(R_1) - C(R_0) = \beta(L_1) \pmod 3$, which is the appropriate form of the parity condition shown in (2) for this setting.

Finally, if $L \neq L_0$ and $L \neq L_1$, then R_0 and R_1 are subsets of *adjacent* regions in the reduced map $M(\mathscr{L}_*)$ (so a parity condition holds for $M(\mathscr{L}_*)$), and R_0 and R_1 are both inside or both outside $W(L_1)$ (so $\gamma(L_1) = 0$). Hence, $C(R_1) - C(R_0) = \{C_*(R_1) - C_*(R_0)\} \pmod 3$. For this case, the appropriate parity condition of the form (2) for $M(\mathscr{L}_*)$ appears as "$C_*(R_1) - C_*(R_0) = \beta(L_2) \pmod 3$" for some properly chosen half-line L_2. It follows that

$C(R_1) - C(R_0) = \beta(L_2)$ (mod 3), which must be the appropriate parity condition of the form (2) for $M(\mathscr{L})$. □

3. Curvilinear maps. A map generated by a finite number of lines is always two-colorable. This applies to *curved* lines as well as straight. (More generally, a map is two-colorable if its vertices are all of even degree. See Theorem 2.3 of Saaty and Kainen (1977).) Likewise, Theorem 1 of Section 2, for *straight* (half-)lines, extends to the curvilinear context with essentially no additional restrictions: One needs to assume that Condition A holds for *curved* lines and *curved* half-lines. Beyond this, one needs to make certain that the curved (half-)lines in (the finite set) $\mathscr{L}$ behave well enough that the generated map $M(\mathscr{L})$ has a *finite number* of *well-defined* regions. Good behavior can be achieved by restricting curved (half-)lines to a *finite* number of (straight) line segments; this will be adequate for our purposes.

Finally, we will assume hereafter, *without explicit mention*, that no (positive length) segment of a curved (half-)line is shared with another curved (half-)line, nor another portion of itself. Such a segment would be contrary to the spirit of this subject. More important, without this assumption, the non-three-colorable map in Figure 1a would be allowed.

Similarly, Theorem 2 of Section 2 extends to the curvilinear context, but *with additional restrictions*; certain kinds of "unwanted crossings" by (half)-lines must be proscribed. Specifically, the proof of Theorem 2 extends, without modification, to curvilinear maps, satisfying Condition A, which have *no (half)-line crossing itself, or one of its sons*: The "wedges" W(L) are now bounded by *curvilinear* half-lines. Likewise, one can define $\alpha(L)$ and $\beta(L)$ in exactly the same way. Then the coloring formula (1) is applicable in precisely the same way. Note that unwanted crossings can not occur with linear maps.

Since it is always possible (we will see) to "remove" the unwanted crossings of a curvilinear map satisfying Condition A, Theorem 2's extension to the curvilinear setting can be used to establish Theorem 1's extension (Theorem 3 below).

Below, "Condition A" and the words "lines" and "half-lines" should be understood in the curvilinear sense, *without mention*, unless required for clarity or emphasis.

Some simple examples appear in Figures 6a,b,c,d with the individual (half–)lines coded to clarify identities.

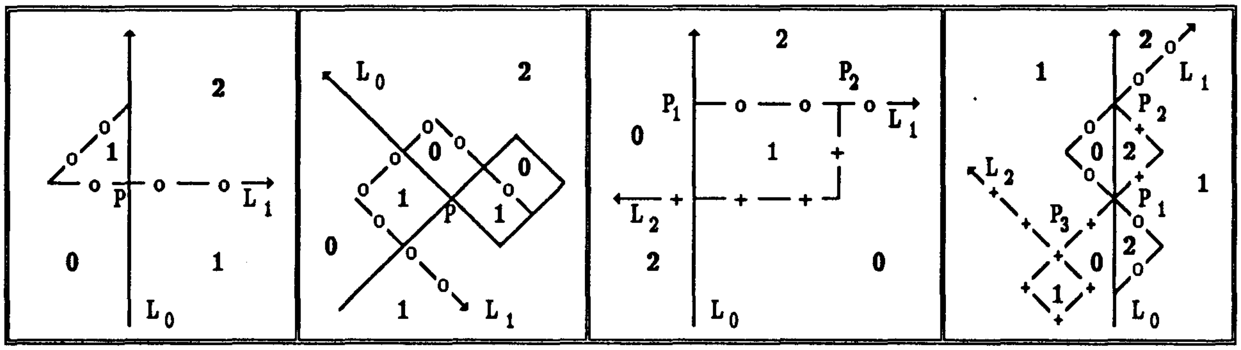

Figure 6a **Figure 6b** **Figure 6c** **Figure 6d**

Each map is three-colorable as shown. The colorings for the first two are unique except for permutations. The region "colored" 2 in Figure 6b is interesting in that it is adjacent to all of the other six regions; so the rest of the map must be two-colored. The coloring formula appearing in (1) applies to the map shown in Figure 6c since it has no "unwanted crossing" (defined above). In particular, (1) yields the colors shown in the figure. Finally, Figure 6d exhibits some of the complexities that can arise with curvilinear maps: Three (half-)lines pass through the vertex at P_1 and two (half-)lines pass through the base point at P_2. The most typical case of a crossing occurs at a vertex of degree four. These arise when one (half–)line crosses another (see point P in Figure 6a), or itself (see point P_3 in Figure 6d). Alternatively, a mere tangency (not shown here), without a "crossing", might occur.

We believe it is instructive to illustrate, with some specific examples, how Theorem 1 can be extended directly to the curvilinear setting. (The approach used here differs somewhat from that used with Theorem 3, which is more easily proven by using our extension of Theorem 2, together with an induction argument that effectively removes unwanted crossings.) Figures 7 and 8 below describe a three-step argument for the maps shown in

Figures 6a and 6b. Notice that the arbitrariness associated with identifying (half-)lines in these maps is actually exploited.

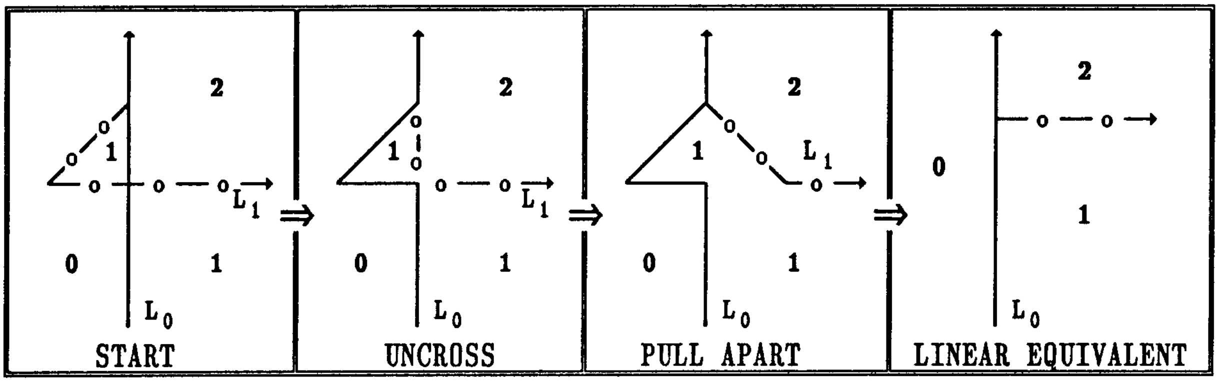

Figure 7

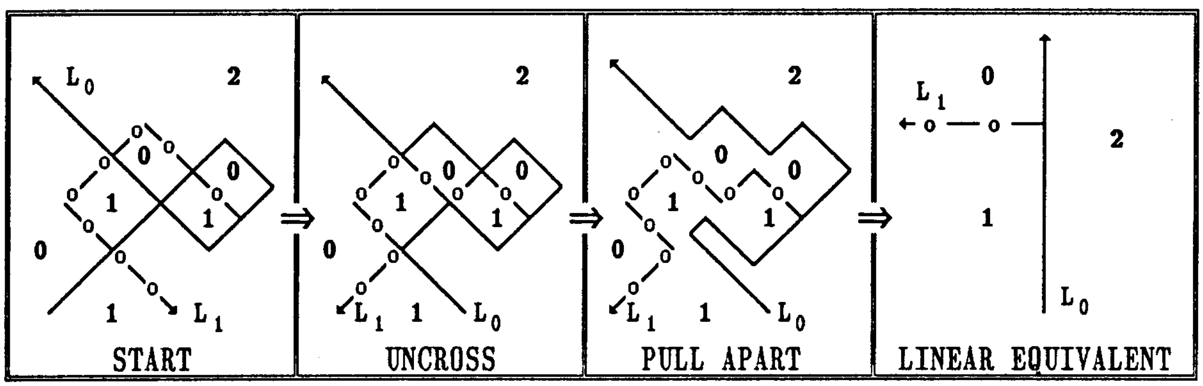

Figure 8

In step 1, one "uncrosses" the crossed (half–)lines by interchanging the "component parts" (parts between vertices) of the (half–)lines. Where (half–)lines formerly crossed (typically at vertices of degree four), one obtains tangencies. This scrambling process does not alter the map $M(\mathscr{L})$ one wishes to three-color, but it can reverse the direction of the "arrow" for some of the component parts. The newly defined collection of (half–)lines must also satisfy Condition A.

In step 2, the map is perturbed enough to separate its tangent parts without otherwise affecting its topological structure. Base points are left undisturbed, so that Condition A continues to hold. A three-coloring of this perturbed map easily leads to a three-coloring of the original map $M(\mathscr{L})$.

In step 3, the perturbed map is identified with a topologically equivalent linear map, which by Theorem 1 is three-colorable. By working backward, one obtains a three-coloring of the original map $M(\mathscr{L})$.

The results for each of the three steps, and the resultant colorings, can be viewed in the various frames of Figures 7 and 8.

The reader will find it instructive to discover the three steps for the maps shown in Figures 6c and 6d.

We are now ready to state and prove the curvilinear analogue of Theorem 1.

Theorem 3. *If a set of curved lines $\mathscr{L}$ satisfies Condition A, then the map* $M(\mathscr{L})$ *is three-colorable.*

Proof. As noted earlier in this section, we will use the curvilinear extension of Theorem 2, together with an induction argument which effectively removes the "unwanted crossings" (somewhat as crossings are removed in "steps 1 and 2" above) without simplifying the task of three-coloring the original map. Recall that an *unwanted crossing* is a crossing of a (half–)line by itself, or by one of its sons. The interested reader can easily confirm that the proof of Theorem 2 extends, without modifications, to curvilinear maps which are free of unwanted crossings.

As a preliminary, we shall begin with "step 2", described above. If there are any tangencies, without a crossing (such as shown in the second frames of Figures 7 and 8), these should be gently pulled apart (as illustrated, with some exaggeration, in the third frames of Figures 7 and 8). All that should remain are genuine crossings.

We remark that the real effect of pulling two tangent lines apart (slightly) is that two regions that have only a point in common merge. Since these regions are allowed to have the same color, and since merging them simply forces them to have the same color, a proper coloring for the perturbed map leads to a proper coloring for the original map.

An induction argument based on the *total* number of crossings will be given. Multi ple crossings will be counted according to their multiplicity. Thus the point P_1 of Figure 6d contributes three crossings. Base points, per se, do not contribute crossings. But crossings can occur at a base point, as at P_2 in Figure 6d.

Each induction step reduces the *total number of crossings* by one. So all of the *unwanted* crossings must eventually disappear. Since the removal of a given unwanted crossing can be accompanied by the conversion of other (previously acceptable) crossings to the "unwanted" classification, the number of *unwanted* crossings might actually increase during a given induction step.

Self-crossings are easily removed: A (half-)line with a self-crossing has a "loop" joined at some self-crossing point P. By simply reversing the direction of the arrow within the loop, the self-crossing is converted to a mere tangency, which, in turn, can be gently pulled apart. (This can easily be seen by examining the point P is Figure 6b, or the point P_3 in Figure 6d. In contrast to what is illustrated in Figure 8, here and below, we are removing a *single* crossing at a time from the original map.)

Next, consider an unwanted crossing of a father A by his son B at a point P. This crossing may occur at a vertex of degree four (such as at P in Figure 6a) or at a vertex of higher degree (such as at P_1 and P_2 in Figure 6d, of degrees six and five, respectively). The point P divides B into two parts B_1 and B_2, ordered in the direction of B's arrow. Similarly, the crossing point P and the base point Q of the son divide A into three parts A_0, A_1 and A_2, ordered in the direction of A's arrow. There are two cases to be considered, depending on the order, with respect to A's arrow, of the points P and Q. These are illustrated in Figures 9.1 and 9.2 for half-lines (the more difficult situation). In both cases, the crossing at P can be converted to a mere tangency at P simply by trading the finite "intervals" A_1 and B_1. Thus the father A is replaced by the concatenated half-line $A_0|B_1|A_2$, and the son B is replaced by the concatenation $A_1|B_2$. For Case 1, the direction of the arrow stays the same in each of the five parts. For Case 2, a reversal occurs for the intervals A_1 and B_1 (see Figures 9.1 and 9.2). It is easily verified that Condition A holds for the new set of (half-)lines.

The final step, in removing the unwanted crossing at P, is to gently pull the newly formed father and son apart at the point of tangency.

Admittedly, it is more difficult to visualize what needs to happen when the vertex at P is of larger degree than four. The crossings of the father L_0 at P_1 by his two sons L_1 and L_2 in Figure 6d are unwanted crossings corresponding to Case 1 and Case 2, respectively. The second crossing of the father L_0 by his son L_1 at P_2 corresponds to Case 1. The third crossing at P_1 by the two sons L_1 and L_2 is benign; it is *not* an unwanted crossing. (Here we view L_2 as a son of L_0. But if we regard L_2 as a son of L_1, then the roles of unwanted and benign crossings are changed.)

After an unwanted crossing has been converted to a tangency, prepared for gentle separation, one is faced with an additional complication if this occurs at a vertex P of degree greater than four. The other "spokes" emanating from this vertex are topologically involved when the separation occurs: The two tangent (half-)lines, A and B, (prepared for separation) divide the immediate neighborhood of P into four regions, two "between the tangency", and one on either side of the tangency. Now consider another (half-)line C which passes through the point P, formed by two of the "other spokes", C1 and C2. If C1 and C2 are both between the tangency, then C becomes completely separated from A and B when the tangency is pulled apart, *or* it could remain tangent to one of them, depending on how the separation is carried out. Alternatively, if C1 is on one side of the tangency, and C2 on the other, then C maintains crossings with A and B after the tangency is pulled apart. Finally, if one of C1 and C2 is between the tangency with the other on one side of the tangency, then C maintains a crossing with one but not both of A and B. None of these possibilities causes problems.

We remark that the complications brought about by vertices of degree greater than four can be finessed, if one chooses to, simply by perturbing the map before beginning the proof of this theorem. It is fairly obvious this can be done legitimately. □

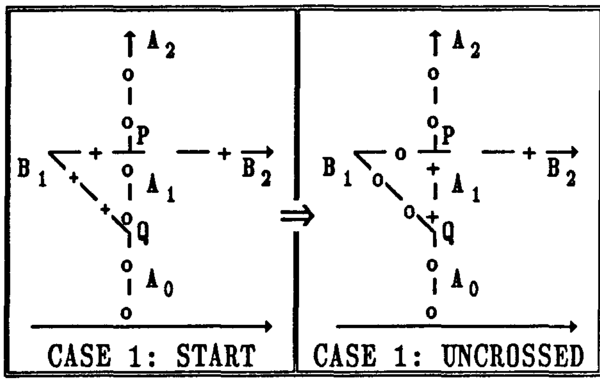

Figure 9.1

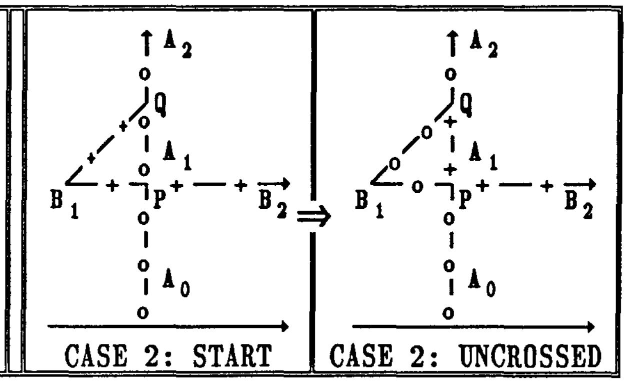

Figure 9.2

References

Appel, K.I. and Haken, W. (1976). "Every planar map is four colorable", *Bull. Am. Math. Soc.,* 82, 711–712.

Brooks, R.L. (1941). "On colouring the nodes of a network", *Proc. Cam. Phil. Soc.,* 37, 194–197.

Garey, M.R., Johnson, D.S. and Stockmeyer, L. (1976). "Some simplified NP-complete graph problems", *Theor. Comput. Sci.* 1, 237–267.

Grötzsch, H. (1958). "Ein Dreifarbensatz fur Dreikreisfreie Netze auf der Kugel", *Wiss. Z. Martin–Luther Univ. Halle–Wittenberg Math. Naturwiss. Reihe,* 8, 109–120.

Grünbaum, B. (1963). "Grötzsch's theorem on 3–colorings", *Mich. Math. J.,* 10, 303–310.

Ore, O. (1967). *The Four–Color Problem,* Academic Press, New York.

Saaty, T.L. and Kainen, P.C. (1977). *The Four Color Problem,* McGraw–Hill, New York.

Stockmeyer, L. (1973). "Planar 3–colorability is polynomial complete", *SIGACT News,* 5, 19–25.