EVERY PLANAR MAP IS FOUR COLORABLE
PART I: DISCHARGING[1]

BY

K. Appel and W. Haken

1. Introduction

We begin by describing, in chronological order, the earlier results which led to the work of this paper. The proof of the Four Color Theorem requires the results of Sections 2 and 3 of this paper and the reducibility results of Part II. Sections 4 and 5 will be devoted to an attempt to explain the difficulties of the Four Color Problem and the unusual nature of the proof.

The first published attempt to prove the Four Color Theorem was made by A. B. Kempe [19] in 1879. Kempe proved that the problem can be restricted to the consideration of "normal planar maps" in which all faces are simply connected polygons, precisely three of which meet at each node. For such maps, he derived from Euler's formula, the equation

$$(1.1) \qquad 4p_2 + 3p_3 + 2p_4 + p_5 = \sum_{k=7}^{k_{max}} (k - 6)p_k + 12$$

where p_i is the number of polygons with precisely i neighbors and k_{max} is the largest value of i which occurs in the map. This equation immediately implies that every normal planar map contains polygons with fewer than six neighbors.

In order to prove the Four Color Theorem by induction on the number p of polygons in the map ($p = \sum p_i$), Kempe assumed that every normal planar map with $p \leq r$ is four colorable and considered a normal planar map M_{r+1} with $r + 1$ polygons. He distinguished the four cases that M_{r+1} contained a polygon P_2 with two neighbors, or a triangle P_3, or a quadrilateral P_4, or a pentagon P_5; at least one of these cases must apply by (1.1). In each case he

produced a map M_r with r polygons by erasing from M_{r+1} one edge in the boundary of an appropriate P_k. By the induction hypothesis, M_r admits a four coloring, say c_r, and Kempe attempted to derive a four coloring c_{r+1} of M_{r+1} from c_r. This task was very easy in the cases of P_2 and P_3. To treat the cases of P_4 and P_5, Kempe invented the method of interchanging the colors in a maximal connected part which was colored by c_r with a certain pair of colors (two-colored chains were later called Kempe chains) to obtain a coloring c_r' of M_r from which one can then obtain a four coloring c_{r+1} of M_{r+1}.

While Kempe's argument was correctly applied to the case of P_4, it was incorrectly applied to the case of P_5 as was shown by Heawood [18] in 1890. Kempe's argument proved, however, that five colors suffice for coloring planar maps and that a minimal counter-example to the Four Color Conjecture (minimal with respect to the number p of polygons in the map) could not contain any two-sided polygons, triangles, or quadrilaterals. This restricts the Four Color Problem to the consideration of normal planar maps in which each polygon has at least five neighbors. Each such map must contain at least twelve pentagons since in (1.1) we have $p_2 = p_3 = p_4 = 0$ and thus

$$(1.2) \qquad\qquad p_5 = \sum_{k=7}^{k_{max}} (k - 6)p_k + 12.$$

Since 1890 a great many attempts have been made to find a proof of the Four Color Theorem. We distinguish two types of such attempts: (i) attempts to repair the flaw in Kempe's work; and (ii) attempts to find new and different approaches to the problem. Among attempts of type (i) we distinguish two subtypes: (i)(a) attempts to find an essentially stronger chain argument for "reducing the pentagon," i.e., proving that a minimal counter-example to the Four Color Conjecture cannot contain any pentagon, and thus does not exist; and (i)(b) attempts to make more extended use of Kempe's arguments in different directions and, instead of "reducing" the pentagon directly, to replace it by configurations of several polygons. Since the method used in this paper is of type (i)(b) we shall restrict our attention to further developments in this branch.

In 1904, Wernicke [28] proved that any normal planar map with $p_2 = p_3 = p_4 = 0$ must contain at least one pentagon which is adjacent to another pentagon or to a hexagon. This result was improved in 1922 by Franklin [14] who proved that either two adjacent pentagons or a pentagon adjacent to two hexagons must occur. A further improvement was made by Lebesgue [21] in 1940 when he displayed a large collection of configurations at least one of which must occur in each normal planar map with $p_2 = p_3 = p_4 = 0$. We refer to such sets of configurations as "unavoidable sets".

In 1913, G. D. Birkhoff [10] used Kempe's chain arguments to develop a general method of proving the "reducibility" of certain configurations, i.e., of proving that these configurations cannot occur in a minimal counter-example

to the Four Color Conjecture. First, he observed that Kempe's work immediately implies the reducibility of any ring (annulus) which is formed by two or three polygons, and that the reducibility of the 4-ring (an annulus formed by four polygons) can be easily proved by Kempe's methods. Then he proved that a 5-ring is reducible provided that there are at least two polygons of the map inside the ring and at least two outside. The corresponding treatment of 6-rings turned out to be much more difficult and was completed much later (1947) by Arthur Bernhart [6]. Birkhoff, however, proved that a 6-ring is reducible in the case that either its interior or its exterior consists of four pentagons. He also proved the reducibility of any polygon which is entirely surrounded by pentagons or by an even number of hexagons.

Regarding the Four Color Problem in general, Birkhoff stated three possible alternatives:

(1) the Four Color Conjecture may be false;
(2) it may be possible to find a collection of reducible configurations such that every planar map must contain one of them (which would prove the Four Color Conjecture);
(3) the Four Color Conjecture may be true but more complicated methods might be required for a proof.

He did not comment, however, on the likelihood he associated with any one of these alternatives.

The methods for proving reducibility described in general by Birkhoff in [10] have since been applied by many investigators, in particular Franklin [14], Errera [13], Winn [29], Chojnacki [11], Arthur Bernhart [7], Heesch [16] and [17], Ore and Stemple [23], Frank Bernhart [8], Allaire and Swart [2], Mayer [22], and Allaire [1]. More detailed descriptions of the algorithms have been given by Heesch [16] and by Tutte and Whitney [27]. In most of the more recent applications [12], [17], [1], [2], [9], [20] the algorithms were carried out by electronic computers. The smallest reducible configurations obtained by these methods each consist of four polygons which surround an edge and have been found by Birkhoff (as described above), Franklin, and Arthur Bernhart.

Regarding the three alternatives given by Birkhoff, the opinions of investigators diverged considerably. Favoring the first alternative (falsity of the Four Color Conjecture), E. F. Moore developed a method of constructing maps which do not contain any reducible configurations which were published prior to the late 1960s. In particular, in March 1977 he constructed a map of 846 polygons[2] which contains no reducible configuration of ring size eleven or smaller (it contains reducible 12-ring configurations, however). This result shows that Birkhoff's second alternative (provability of the Four Color

_____

[2] When this paper was submitted it was thought that a map of 341 polygons, constructed by Moore in January 1963, had this property. But J. Graber found a reducible 11-ring configuration in that map.

Conjecture by means of an unavoidable set of reducible configurations) cannot lead to a short proof (and not even to a moderately long one) since such a proof would require at least reducible 12-ring configurations and very likely a large number of these.

On the other hand, Heesch [16] observed that one finds large numbers of reducible configurations if one considers configurations which are not too small in size and contain relatively many pentagons. He stated a more detailed version of Birkhoff's second alternative as a conjecture [16; p. 216] by conjecturing that the reducible configurations in the finite (unavoidable) set will range in size up to the second neighborhood of a pair of polygons. Heesch states his results in the dual language of planar triangulations and vertex colorations and uses a special coding for indicating the degrees of the vertices (which we use in this paper also; see Figure 1 of Section 2). Heesch stated his conjecture in a colloquium talk which he gave at the University of Kiel (Germany) in about 1950.

In [16] Heesch treats several special cases of triangulations and proves that each of them contains a reducible configuration. The case of triangulations without vertices of degrees six or seven (which had earlier been taken care of by Chojnacki [11] by a different method) is treated by a method which we call a discharging procedure. Each degree-5-vertex is regarded as carrying a positive "charge" of 60 and each vertex of degree $k \geq 7$ is regarded as carrying a negative "charge" of 60 $(6 - k)$. Then by (1.2), the sum of the charges is positive (720). In a "first discharging step" the positive charges of the 5-vertices are distributed in equal fractions to their major neighbors (where major means of degree $k \geq 7$). Then it is shown that positive charges can occur only in 16 special cases provided that the triangulation does not contain one of a list of 20 reducible configurations. Each of the 16 special cases is described by a configuration in which a major vertex has received so much positive charge from its degree-5-neighbors that its charge (which was initially negative) has a value $z > 0$. These configurations are called z-positive configurations and are not reducible (nor do they contain any reducible sub-configurations). Then in a "second discharging step" the new positive charges z are distributed to currently negative neighbors and it is shown that no positive charges remain, of course provided that none of the 20 reducible configurations is contained in the triangulation. This implies that there does not exist any triangulation without 6- and 7-vertices which does not contain at least one of the 20 reducible configurations (since the sum of all charges must be positive).

Haken, who had been a student at Kiel when Heesch gave his talk, communicated with Heesch in 1967 inquiring about the technical difficulties of the project of proving Heesch's conjecture and the possible use of more powerful electronic computers.

In 1970 Heesch communicated to Haken an unpublished result which he later referred to as a finitization of the Four Color Problem, namely that the first discharging step (described above), if applied to the general case, yields about

8900 z-positive configurations (most of them not containing any reducible configurations) which he explicitly exhibited. He hoped that it would be possible to find a correspondingly large set of reducible configurations and to work out a second discharging step in order to obtain an unavoidable set consisting only of reducible configurations and thus a proof of the Four Color Conjecture. Haken was very pessimistic regarding the combinatorial complexity of this task. He proposed to search for a better discharging procedure in order to reduce the complexity of the project. Returning to the special case of triangulations without 6- and 7-vertices, Haken immediately found an improvement of the discharging procedure which shortened the treatment of the case considerably [15]. Encouraged by this result, Haken made several suggestions for developing an equally improved discharging procedure for the general case.

Heesch asked Haken to cooperate on the project and, in 1971, communicated to him several unpublished results on reducible configurations, in particular, his observation of three "reduction obstacles," called four-legger vertices (i.e., vertices in the configuration with four or more neighbors in the ring surrounding the configuration), three-legger articulation vertices (a vertex the removal of which disconnects the configuration is called an articulation vertex), and hanging 5-5-pairs (pairs of adjacent 5-vertices connected by edges to only one other vertex of the configuration). The presence of one of these obstacles appears to prevent the reducibility of the configuration (unless the configuration contains a proper sub-configuration which is reducible and does not contain the obstacle).

The cooperation between Heesch and Haken was interrupted in October 1971 when the work of Shimamoto was thought to have settled the Four Color Problem. Actually Shimamoto's work stimulated Tutte and Whitney to work out the first published theory of reduction obstacles in [27]. Their methods were subsequently used by Stromquist [26] for exhibiting several further reduction obstacles which included the three types mentioned above as the practically most important special cases.

Early in 1972 Haken proposed, as a first step towards a proof of Heesch's conjecture, to develop a discharging procedure which would yield an unavoidable set of configurations which do not contain the first two reduction obstacles mentioned above. Such configurations were called geographically good. The object was to shift the emphasis from computing of reducible configurations to improving the discharging procedure and to obtain a more reliable estimate of the number and the size of the reducible configurations which would eventually be required for the proof. An investigation in this direction was immediately made by Osgood [24], who treated the special case of triangulations in which every vertex has degree five, six, or eight.

The complexity of Osgood's work convinced Haken that, even with a further improved discharging procedure, the general case could not be effectively attacked without the aid of electronic computers for the tedious task of discussing all possible configurations which can be obtained by merging pairs, triplets,

etc., of given configurations. It is obvious that this task arises again and again if one must survey all cases in which two, three, or more dischargings "overcharge" a major vertex (i.e., make a formerly negative vertex positive).

At this time, in May 1972, Appel suggested proceeding with the project since he felt that the necessary computer work was quite feasible.

From 1972 through 1975 the two authors gradually improved their discharging procedure. A discharging procedure may have three major types of defects: (i) In some situations a positive charge may remain without a reducible (or at least likely to be reducible) configuration occurring; (ii) The number of essentially different situations in which a positive charge remains is excessive (in this context, two situations are "essentially different" if they do not yield the same reducible configuration); (iii) In some situation in which positive charge remains, even the smallest reducible (or likely to be reducible) configuration which occurs is of excessive ring size.

The search for major defects is usually very tedious by hand (unless the defects are very obvious) but can be done quite effectively by a computer which enumerates certain situations of remaining positive charge and tests each of them for acceptable sub-configurations (i.e., likely to be reducible, or at least geographically good configurations of limited ring size). An example of such a computer program has been described in more detail in [4, Section 27]. When a major defect was found, the authors had to find a way of changing the discharging procedure to avoid the defect. Then, usually the computer program had to be changed accordingly and the search for further defects could begin.

In 1974, the authors could prove the existence of a finite unavoidable set of geographically good configurations [4] and describe an algorithm for constructing such a set. A much shorter existence proof (but without a construction algorithm) was found shortly later by Stromquist [26, Chapter 4]. The discharging procedure of [4] was rather complicated and was illustrated later by applying it to the special case of "isolated 5-vertices" in [3]; its application to the general case was never fully worked out because many more improvements were found during 1975.

The cooperation with Heesch did not resume after 1972, however, and no agreement could be reached as to which method of attacking the general case would be better.

By September 1975 the authors had improved the methods so far that it seemed to be more work to change the computer program after each new improvement than to carry out the case enumeration by hand. It was decided at this point that the most efficient approach would be to proceed with the discharging procedure by hand and shift the computer work to the computation of reducible configurations as described in Part II. (Actually, the authors and John Koch had begun to study reducibility algorithms for computers in late 1974. Initially they conjectured that different algorithms from the standard ones would be required for the great number of configurations of large size which

were expected, but the efficiency of the discharging procedure rendered this unnecessary.)

In January 1976 Jean Mayer found a considerable improvement of the procedure for treating the special case of isolated 5-vertices. A few weeks earlier the authors had found a corresponding improvement of their procedure for the general case. It appeared interesting to compare different discharging procedures by applying them to the same special case. Therefore the authors applied their discharging procedure (essentially the same one as used in this paper) to the case of isolated 5-vertices and accepted an invitation of Mayer to write a joint paper [5]. Mayer's procedure is still considerably simpler and more effective than that of the authors', but, thus far, no corresponding simplification for the general case has been found. The authors are aware, however, of certain possible improvements in different directions; these are discussed in Section 5 of this paper.

Other interesting comparisons between different discharging procedures are possible by examining the works of Ore and Stemple [23], Stromquist [26, Appendix], and Mayer [22] on the Birkhoff number (treating triangulations with an upper bound on the number of vertices) and of Stanik [25] and Allaire [1] (on triangulations without 6-vertices).

2. The discharging procedure

We consider a triangulation Δ of a closed 2-dimensional manifold M^2 of Euler characteristic χ. We assume that Δ is a simplicial complex, i.e., that it does not contain any loops, nor any 2-circuits, nor any 3-circuits other than boundaries of triangles. Moreover we assume that every vertex V of Δ satisfies deg $(V) \geq 5$ (i.e., has degree at least 5).

To every vertex V of Δ we assign an *initial charge* $q_0(V) = 60\,(6 - \deg\,(V))$. Then, by Kempe's version of Euler's formula, we have

$$(2.1) \qquad\qquad \sum_{V \in \Delta} q_0(V) = 360\chi.$$

Note that $q_0(V_k)$ is negative for all major vertices ($k \geq 7$), zero for all 6-vertices, and positive for all 5-vertices.

We shall describe *configurations* C mainly by drawings (of bounded, planar, connected, simply connected triangulations) in which *degree specifications* of the vertices are indicated by the symbols introduced by Heesch [16] (see Figure 1). A configuration C is said to be *contained in* Δ if there is a simplicial immersion $f: C \to \Delta$ which respects the degree-specifications. Here we use the following definitions. (For a more elaborate treatment of these basic concepts see Section 8 of [4].)

We call a continuous mapping $f: C \to \Delta$ of a configuration C into a triangulation Δ an *immersion* (which respects the degree specifications) if it has the following three properties.

●	degree 5		degree ≳ 6
⟨	(no special marking at the vertex) degree 6	∩	degree ≳ 7
○	degree 7		not both degrees 6 (but each ≳ 6)
◻	degree 8	∏	degree ≳ 8
▽	degree 9		degree unspecified (but ≳ 5)
◇	degree 10	⟨≳10⟩	degree ≳ 10
	degree 11	m	degree 5 or 6 ("minor")
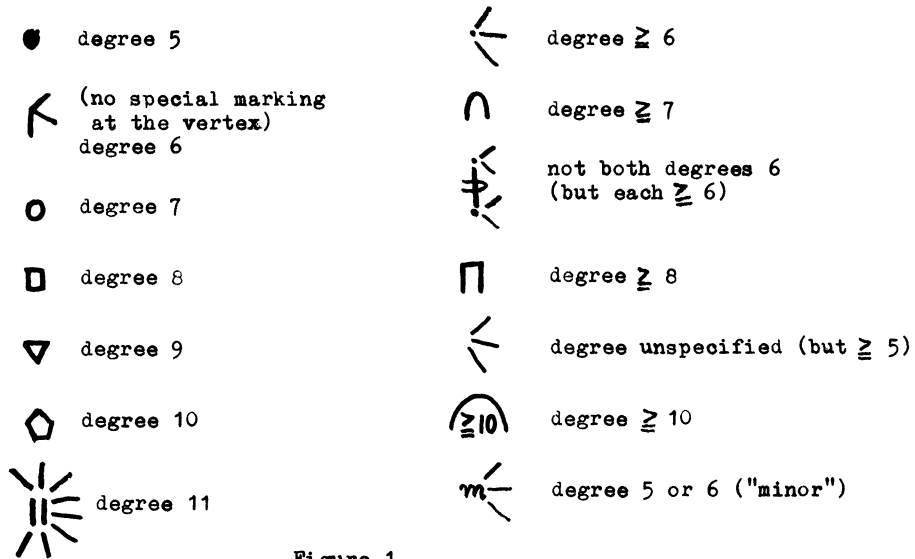
Figure 1

(i) *f* is *simplicial* and *dimension-preserving*; i.e., if σ is a simplex (vertex, edge, or triangle, respectively) of *C* then *f* | σ is a homeomorphism onto a simplex σ' of Δ.

(ii) *f respects the degree specifications* of *C*; i.e., if *V* is a vertex of *C* for which a degree is completely or partially specified (e.g., deg (*V*) = 6 or deg (*V*) ≥ 7) then the degree of *f*(*V*) in Δ agrees with this specification.

(iii) There is a small neighborhood *N* of *C* in the plane so that *N* is a disk (while *C* may have articulation points) and there is an extension $\tilde{f}$ of *f* over *N* so that $\tilde{f}$ is *locally one-to-one*; i.e., if *p* is a point of *N* and *U* is a small neighborhood of *p* in *N* then $\tilde{f}$ | *U* is a homeomorphism.

If a configuration *C* is said to be *contained in* another configuration, *D*, then this means that there is a simplicial immersion *f*: *C* → *D* which respects the degree specifications. However, in all cases which we have to consider explicitly in this paper, the configurations *C* and *D* will be so small that the immersion *f* will be an *embedding* (i.e., one-to-one and not only locally one-to-one) and *C* will be said to be *contained in D as a sub-configuration*. In general, the immersion or embedding *f* may preserve or reverse the orientation of *C*; in particular, if *C* and *D* are defined by drawings in the plane we may or may not wish to distinguish the cases that *C* is contained *reflected* or *non-reflected* in *D*. If we want to indicate that *C* is contained non-reflected in *D* we write "*Cn is contained in D*"; if we want to say that *C* is contained reflected in *D* (i.e., that the mirror image of *C* is contained in *D*) then we write "*Cr is contained in D*". The expression "*C is contained in D*" means that *Cn or Cr* is contained in *D*.

Now we shall define[3] a *discharging procedure* $\mathscr{P}$ which can be applied to any triangulation Δ with the above properties and which assigns to each vertex V of Δ a *terminal charge* $q(V)$ so that again

(2.2) $$\sum_{V \in \Delta} q(V) = 360\chi.$$

We define the charge function q on (the set of vertices of) Δ by deriving it from q_0 by transfers of charge from 5-vertices V_5 to nearby major vertices. We distinguish two kinds of such charge transfers: (i) short range dischargings which transfer charge along edges of Δ which join V_5's to major vertices, and (ii) transversal dischargings which transfer charge from a V_5 across one, two, or three 6-6 edges (edges joining pairs of 6-vertices) to a major vertex; these are abbreviated T-*dischargings*. First, we define the T-dischargings. If one of the

T1 #1   T1 #2   T1 #3   T1 #4

T2 #5   T2 #6   T2 #7
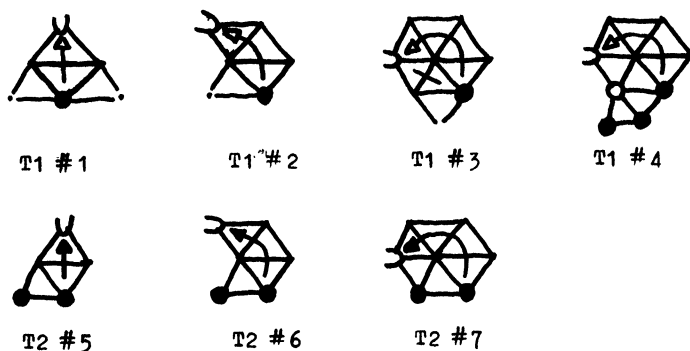
Figure 2

seven configurations of Figure 2, which we refer to as T-*discharging situations*, is contained in Δ then charge is transferred as indicated by the arrows. The

solid arrow means a transfer of 20, the open arrow means a transfer of 10, with the following exceptions (see Figure 3) in the case that two T-dischargings leave the same V_5 across the same 6-6 edge (but arrive at different major vertices). If at least one of the two arrows is solid then 10 is transferred along each of the arrows; if both of the arrows are open then 5 is transferred along each of

---

[3] See discussion of choice of discharging procedure in Section 5 (p. 487) for a general description of the method of defining the discharging procedure.

them. In this case we may use the symbol consisting of a solid arrow splitting into two open arrows or an open arrow splitting into two skeletal arrows of value 5 as in Figure 3.
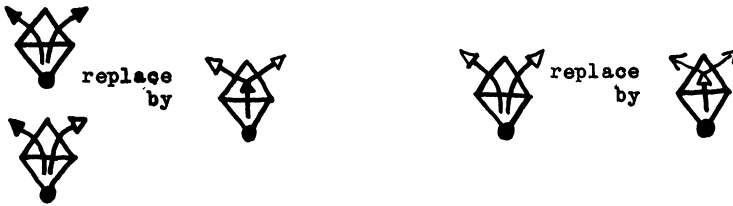
Figure 3

We refer to a T-discharging arrow of value 5 or 10 as a T1-*discharging* and to an arrow of value 20 as a T2-*discharging*. (Correspondingly we have named the seven T-situations in Figure 2 as T1#1,..., T1#4, T2#5,..., T2#7.) Occasionally we draw an "arrow without head" to indicate the path of a T-discharging whose value we do not choose to specify. For instance, in the configuration T1#3 (Figure 2) the unspecified T-discharging will be T2 if the vertex without degree-specification (at the bottom of the drawing) is mapped to a V_5 (of Δ) and will be T1 otherwise.

In order to define the short-range dischargings, let E be an edge of Δ which joins a 5-vertex V_5 to a major vertex V_k ($k \geq 7$). In Table 1 we have defined (by individual drawings) the "situations of small dischargings," abbreviated S-*situations*. Most of the configurations drawn in Table 1 contain some vertices of partially specified degrees (≥ 6 or ≥ 7) which are separated from the configuration by "clip marks."
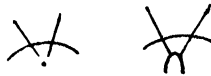
The configurations without these "clipped off" vertices are the S-situations. The full configurations are called *enlarged S-situations* and will be explained later (Lemma S^+ in Section 3). Each of these configurations has a *distinguished edge* which is drawn vertical and marked by a number (0, 5, 10, 15, 20, or 25), its *discharging value*. We say that an S-situation C *applies at E* if C is contained in Δ in such a way that the distinguished edge D of C is identified to E. The "situations of large discharging," called L-*situations*, have been similarly defined by drawings in Table 2. Again, the discharging values (35, 40, 50, or 60) are marked at the distinguished edges.

In the drawings we have used the following special abbreviations, which are explained by examples in Figure 4.

means attachment of some T2-situation so as to induce a T-discharging of 20 to the major vertex as indicated by the arrow.

means attachment of some T-situation (T1 or T2) so as to induce a T-discharging as indicated by the arrow.

means a major vertex V, excepting one disposition: V cannot be a V_7 which is adjacent to a V_5 which itself does not belong to the configuration but is adjacent to another vertex of the configuration.

means a major vertex V, excepting the disposition in which the vertex V occurs as in the partial diagram at right.

means that *no* T-discharging crosses the marked 6-6 edge.

Regarding the edge E we now have three possibilities (see Figure 5 for examples):

Case (i).   No S- or L-situation is attached at E. In this case we call E a *regular discharging edge* or R-*edge* and we define the discharging value $d(E)$ of E to be 30.

Case (ii).   One or more S-situations, but no L-situations, are attached at E Then we call E a *small discharging edge* or S-*edge* and we define its discharging value $d(E)$ to be equal to the smallest of the discharging values of the attached S-situations.

Case (iii).   One or more L-situations and zero, one, or more S-situations are attached at E. Then we call E a *large discharging edge* or *L-edge* and we define $d(E)$ to be equal to the *largest* of the discharging values of the attached L-situations.

Now every edge E of Δ which joins a V_5 to a major vertex has a uniquely defined discharging value $d(E)$. (To all other edges we may assign the discharging value zero.) We obtain the charge distribution q from q_0 by (simultaneously) transferring along each edge E the charge $d(E)$ (from the V_5 to the major vertex) and carrying out the T-dischargings. We refer to the charge transfers along R-, S-, and L-edges as R-, S-, or L-*dischargings*, respectively. This finishes the definition of our discharging procedure $\mathscr{P}$.

stands for one of the 5 configurations below,
    provided that the T2-discharging arrows do not split,
    i.e., in particular, that deg(A)≤ 6 in the last three
    configurations
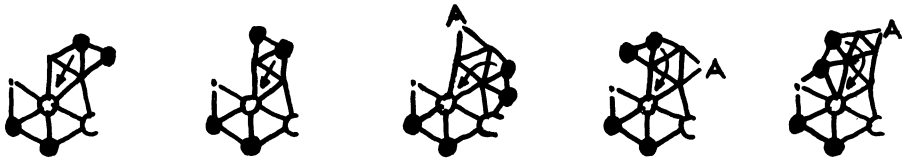
stands for one of the 10 configurations above and below,
    where in this case all T-discharging arrows may or may not
    split (i.e., in particular, deg(A) is arbitrary above)
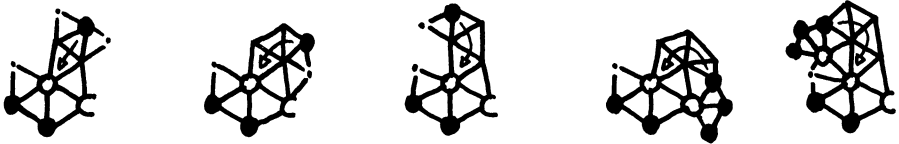
excludes   and

excludes

Figure 4

20 # 001   10 # 002   10 # 003   20 # 004   20 # 005

20 # 006   20 # 007   20 # 008

5 # 011   0 # 012   0 # 013   0 # 014   0 # 015

0 # 016   0 # 017

10 # 021   10 # 022   10 # 023   10 # 024   10 # 025

10 # 026   10 # 027   10 # 028   10 # 029   10 # 030

10 # 031   10 # 032   10 # 033   10 # 034   10 # 035

Table 1, page 1

10 # 036   10 # 037   10 # 038   10 # 039   10 # 040

10 # 041

20 # 051   20 # 052   20 # 053   20 # 054   20 # 055

20 # 056

20 # 061   20 # 062   20 # 063   20 # 064   20 # 065

20 # 066   20 # 067   20 # 068   20 # 069   20 # 070

20 # 071   20 # 072   20 # 073   20 # 074   20 # 075

Table 1, page 2

20 # 076

20 # 079

20 # 080

20 # 081

20 # 082

20 # 083

20 # 084

20 # 085

20 # 086

20 # 087

20 # 088

20 # 089

20 # 090

20 # 091

25 # 092

20 # 093

25 # 094

20 # 095

20 # 096

20 # 097

25 # 098

20 # 099

20 # 100

25 # 101

20 # 102

20 # 103

20 # 104

20 # 105

20 # 106

20 # 107

20 # 108

20 # 109

20 # 110

Table 1, page 3

20 # 111   20 # 112   20 # 113   20 # 114   20 # 115

20 # 116   20 # 117   20 # 118   20 # 119   20 # 120

20 # 121   20 # 122   20 # 123   20 # 124   20 # 125

20 # 126

20 # 131   20 # 132   20 # 133   20 # 134   20 # 135

20 # 136   20 # 137   25 # 138   20 # 139   20 # 140

20 # 141   20 # 142   20 # 143   25 # 144   25 # 145

Table 1, page 4

25 # 146   20 # 147   20 # 148   20 # 149   20 # 150

0 # 151

10 # 161   10 # 162   10 # 163   10 # 164   10 # 165

10 # 166   10 # 167

20 # 171   20 # 172   20 # 173   20 # 174   20 # 175

20 # 176   20 # 177   20 # 178   20 # 179   20 # 180

20 # 181   20 # 182   20 # 183   20 # 184   20 # 185

Table 1, page 5

20 # 186   20 # 187   20 # 188   20 # 189   20 # 190

20 # 191   20 # 192   20 # 193   20 # 194   20 # 195

20 # 196   25 # 197   20 # 198   20 # 199   20 # 200

20 # 201   20 # 202   20 # 203   20 # 204

20 # 211   20 # 212   20 # 213   20 # 214   20 # 215

20 # 216   20 # 217   20 # 218   20 # 219   20 # 220

25 # 221   20 # 222   20 # 223   20 # 224   20 # 225

Table 1, page 6

20 # 226   20 # 227   20 # 228   20 # 229   20 # 230

0 # 231   0 # 232   0 # 233   0 # 234   0 # 235

0 # 236   0 # 237   0 # 238

20 # 230.1

10 # 241   10 # 242   10 # 243   10 # 244   10 # 245

10 # 246   10 # 247   10 # 248   10 # 249   10 # 250

10 # 251   10 # 252   15 # 253

20 # 261   20 # 262   20 # 263   20 # 264   20 # 265

Table 1, page 7
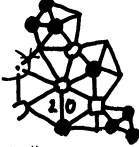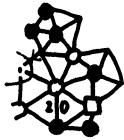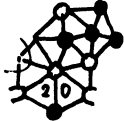
20 # 266   20 # 267   20 # 268   25 # 270

20 # 271   20 # 272   20 # 273   20 # 274   20 # 275

20 # 276   20 # 277   20 # 278   20 # 279   20 # 280

20 # 281   20 # 282   20 # 283   20 # 284   20 # 285

20 # 286   20 # 287   20 # 288   20 # 289   20 # 290

20 # 291

20 # 301   20 # 302   20 # 303   20 # 304   20 # 305

Table 1, page 8
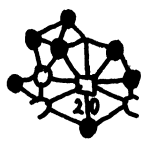
20 # 306   20 # 307   20 # 308   20 # 309   20 # 310

20 # 311   25 # 312   20 # 313   25 # 314   20 # 315
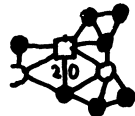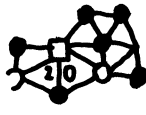
20 # 316   20 # 317   20 # 318   20 # 319   20 # 320

20 # 321   20 # 322   20 # 323   20 # 324   20 # 325

20 # 326   20 # 327   20 # 328   20 # 329

Table 1, page 9

40 #401   40 #402   40 #403

60 #411

40 #421   40 #422   40 #423   40 #424   40 #425

40 #426   40 #427   40 #428

40 #431   40 #432   40 #433   40 #434   40 #435

40 #436   35 #437

50 #441

Table 2, page 1

40 #451   40 #452   40 #453   40 #454   40 #455

40 #456   40 #457   40 #458   40 #459   40 #460

40 #461   40 #462   40 #463   40 #464   40 #465

40 #466   40 #467   40 #468   35 #469

40 #471   40 #472   40 #473   35 #474   40 #475

40 #476   40 #477   35 #478   40 #479   40 #480

35 #481

Table 2, page 2

K. APPEL AND W. HAKEN
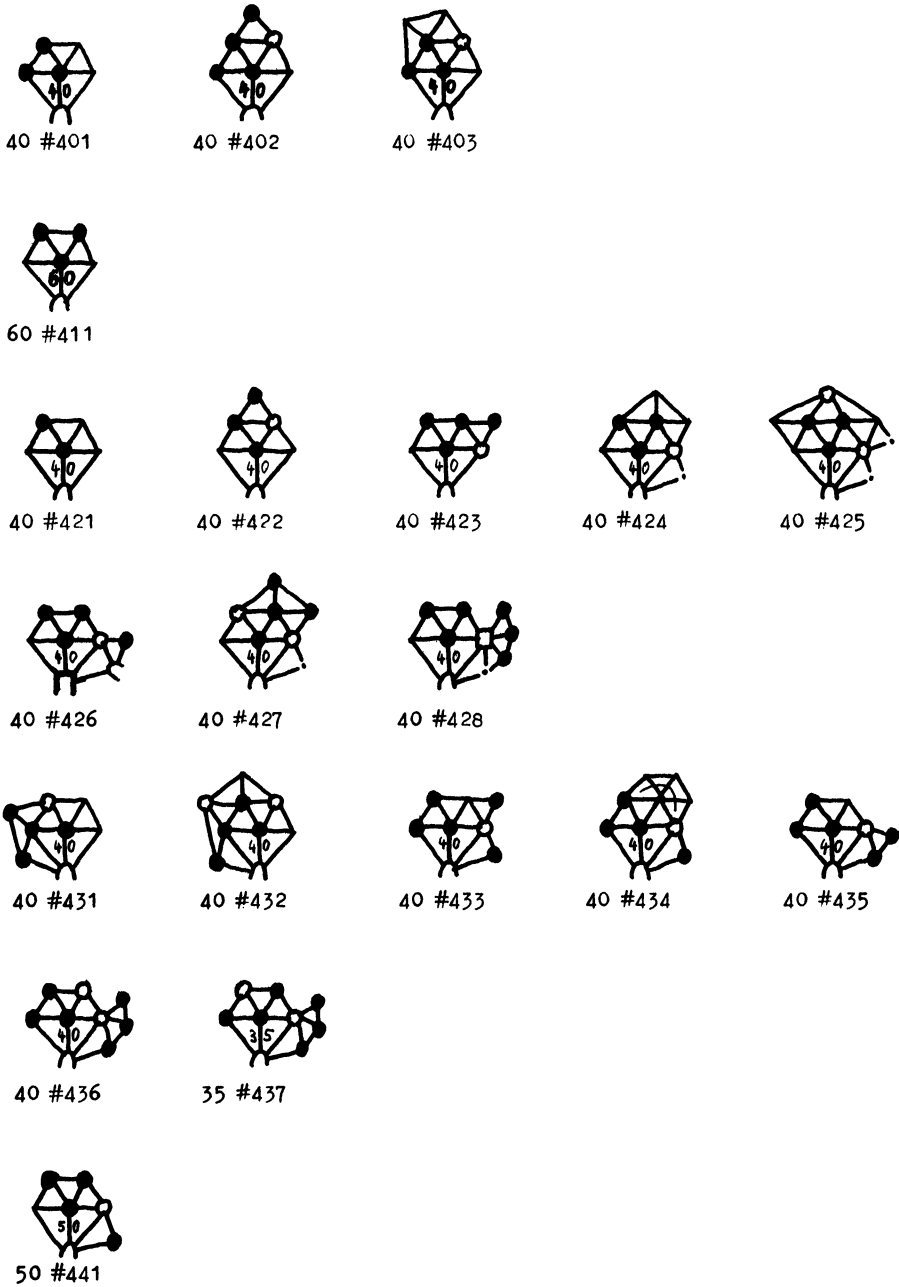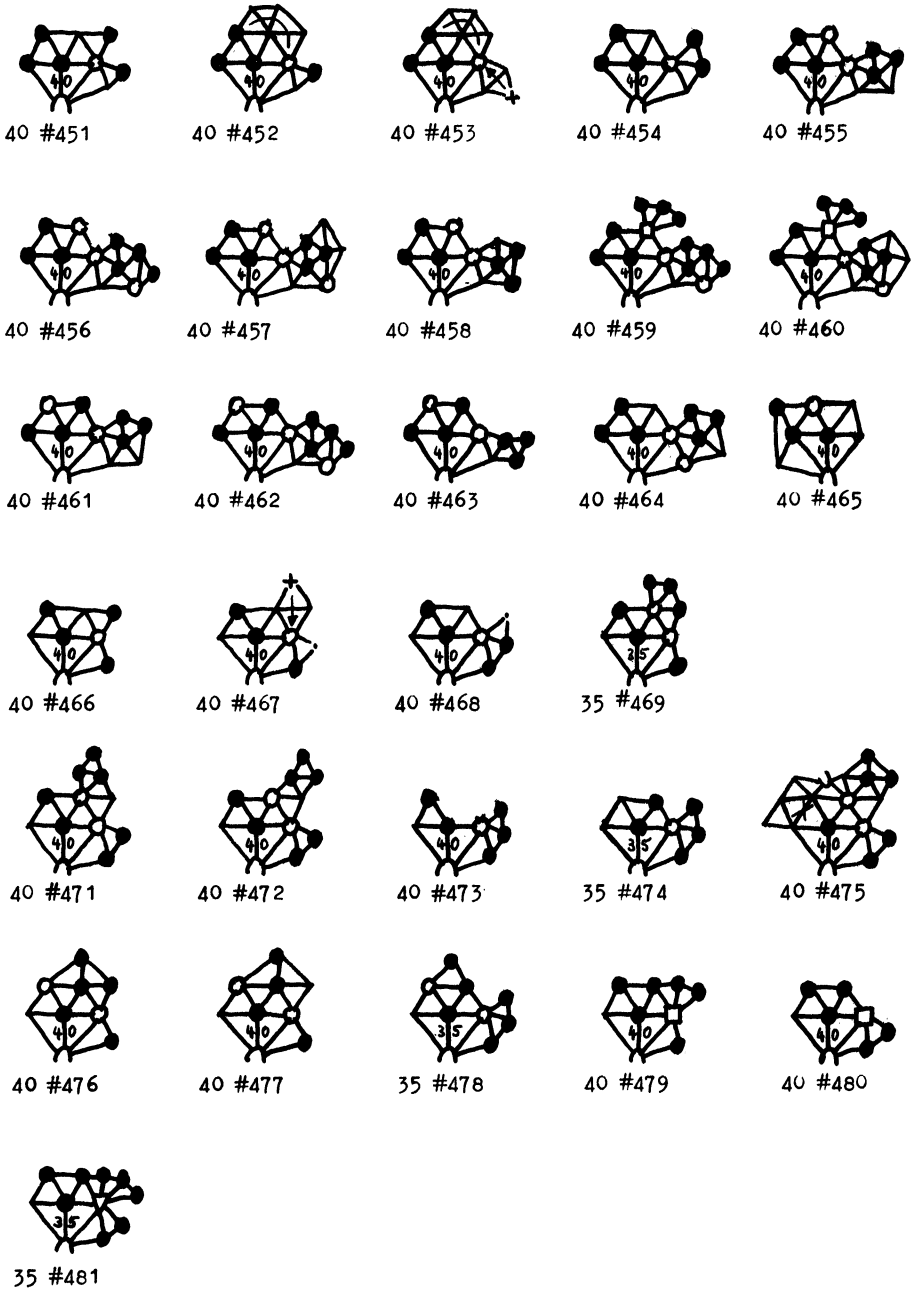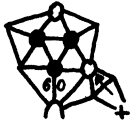
60 #491

50 #492

50 #493

50 #494

50 #495

40 #496

40 #497

40 #498

40 #499

40 #500

35 #501

40 #502

40 #503

40 #505

40 #506

40 #507

40 #508

40 #509

40 #510

40 #511

40 #512

40 #513

40 #514

40 #515

40 #516

40 #517

40 #518

40 #519

40 #520

40 #521

35 #522

40 #530

Table 2, page 3

40 #531   40 #532   35 #533   35 #535

40 #536   40 #537   40 #538   40 #539   40 #540

40 #541   40 #549   40 #550

50 #551   50 #552   50 #553

40 #561   40 #562   35 #563   35 #564   40 #565

40 #566   40 #567   40 #568   40 #569   40 #570

40 #571   40 #572   40 #573   40 #574   40 #575
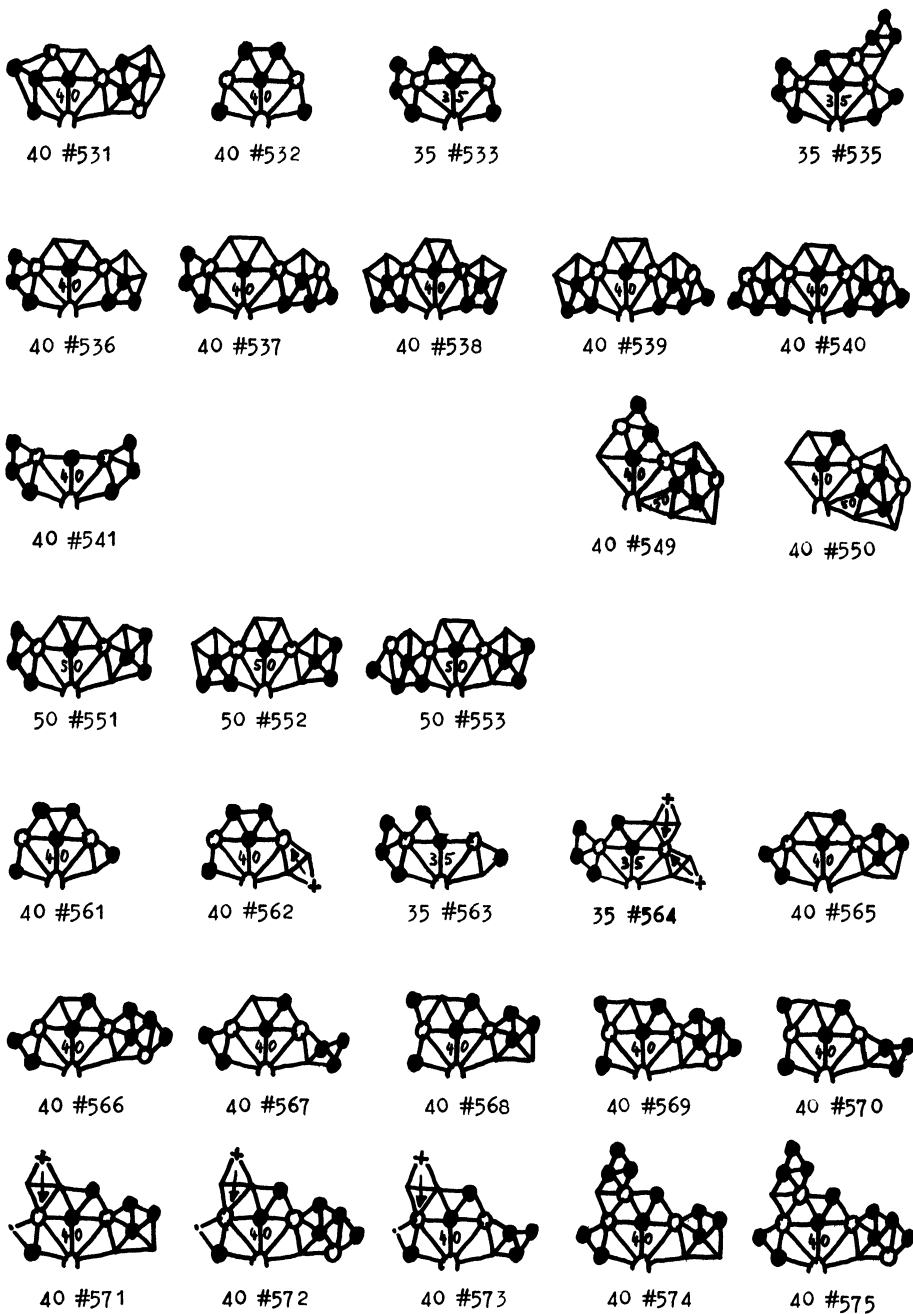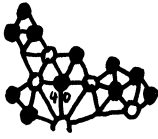
Table 2, page 4
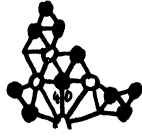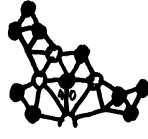
40 #576   40 #577   40 #578   40 #579   35 #580

40 #581   40 #583   40 #584   40 #585

40 #586   40 #587   40 #588   40 #589   40 #590

40 #591   40 #592   40 #595
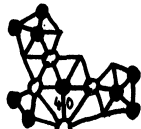
40 #597   40 #598   40 #599   40 #600

40 #601   40 #602   35 #603   40 #604   40 #605

40 #606   40 #607   40 #610

Table 2, page 5

40 #611   40 #612   40 #613   40 #614   40 #615

40 #616   40 #617   40 #618   40 #619   40 #620

60 #621   50 #622   50 #623   50 #624

40 #631   40 #632   40 #633   40 #634   40 #635

40 #637   40 #639

40 #641   40 #643   40 #645
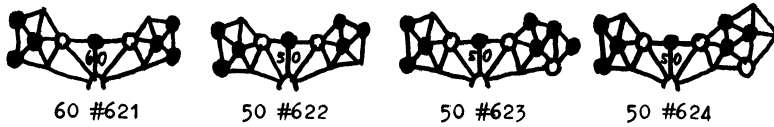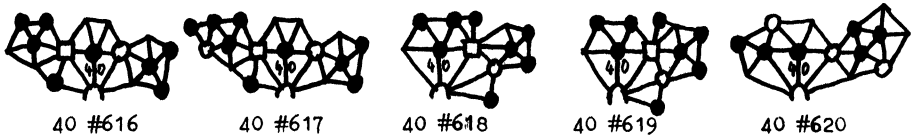
Table 2, page 6

40 #659   40 #660

40 #661   40 #662   40 #663   40 #664   40 #665

40 #668   40 #669   40 #670

40 #671   40 #672   40 #673   40 #674   40 #675

40 #676   40 #677   40 #678   40 #679   40 #680

40 #681   40 #682   40 #683   40 #684   40 #685

Table 2, page 7

40 #686   35 #687   40 #688   40 #689   40 #690

40 #691

40 #701   40 #702

40 #711   40 #712   40 #713   40 #714   35 #715

40 #716   40 #718   50 #720

40 #721   40 #722   40 #723   40 #725

40 #726   40 #727   35 #728   40 #729   40 #730
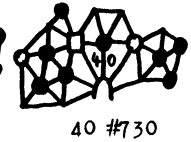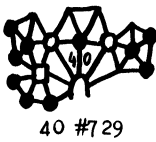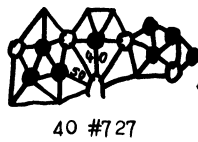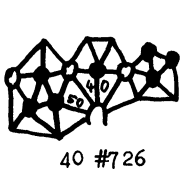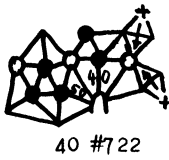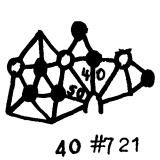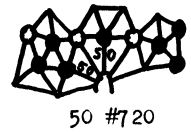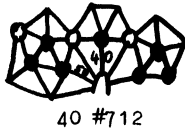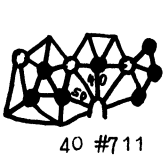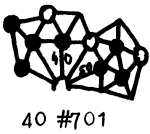
Table 2, page 8
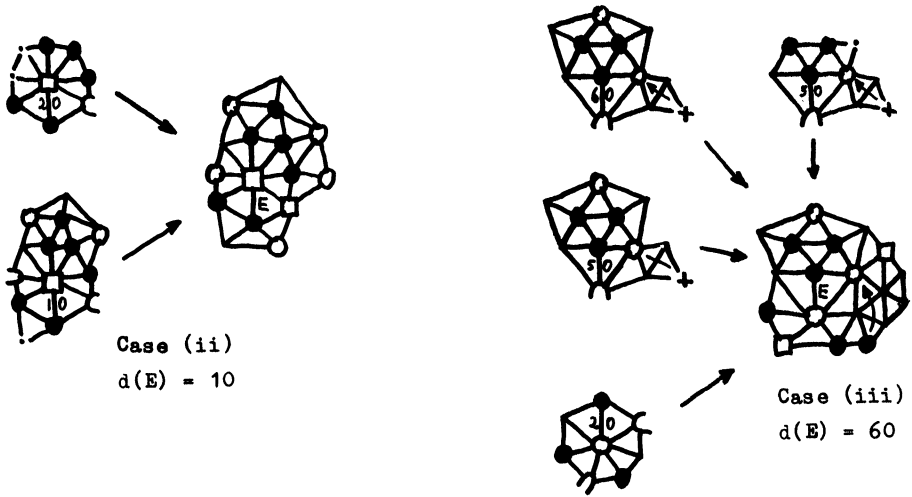
Figure 5

Note that the discharging procedure $\mathscr{P}$ depends essentially on the set $\mathscr{T}$ of T-discharging situations as defined in Figure 2, on the set $\mathscr{S}$ of S-situations, and on the set $\mathscr{L}$ of L-situations as defined in Tables 1 and 2. Thus, to be precise, we should denote our discharging procedure by $\mathscr{P}(\mathscr{T}, \mathscr{S}, \mathscr{L})$, indicating that we would obtain different discharging procedures by using different sets of T-, S-, or L-situations.

We remark that most of the S-situations in $\mathscr{S}$ (197 out of 269) have discharging value 20 and most of the remaining ones have values 0 or 10. Value 0 occurs 15 times (##012, ..., 017, 151, 231, ..., 238) in Table 1; Value 10 occurs 42 times (##002, 003, 021, ..., 041, 161, ..., 167, 241, ..., 252); Value 5 occurs only once (#011); Value 15 occurs only once (#253); and Value 25 occurs 13 times (##092, 094, 098, 101, 138, 144, 145, 146, 197, 221, 270, 312, 314). Correspondingly we distinguish three classes of S-edges for which we use the following abbreviations.

S0 means discharging value 0 or 5,
S1 means discharging value 10 or 15,
S2 means discharging value 20 or 25.

In Table 1 we have ordered the S-situations with respect to the following distinctions:

(1) The degrees of the two vertices adjacent to the distinguished edge (see for instance A and B in 20#001). Type 5-5 occurs only once (#001); ##002, ..., 008 are of Type 5-6; Type 6-6 does not occur at all; ##011, ..., 150 are of

Type 5-major; ##151, . . . , 230.1 are of Type 6-major; and ##231, . . . , 329 are of Type major-major.

(2) The classes S0, S1, S2 of discharging values.

(3) The degrees of the distinguished major vertices (7, 8, and 9 occur; 9 occurs only in Type 5-major).

(4) The degrees of the neighbors of the distinguished major vertex in lexicographic order (reading counterclockwise and starting at the distinguished V_5).

Regarding the L-situations we use the following abbreviations.

L4 means value of the distinguished edge is 40 or 35,
L5 means value of the distinguished edge is 50,
L6 means value of the distinguished edge is 60.

Note that L5 occurs only 12 times (##441, 492, . . . , 495, 551, 552, 553, 622, 623, 624, 720) and L6 occurs only three times (##411, 491, 621).

The ordering of the L-situations in Table 2 has been determined according to the following distinctions:

(a) The *width w* of the L-situation, i.e., the number of fully specified neighbors of the *pivot* (the distinguished major vertex). Situations ##401, . . . , 428 are of Type $W3$ ($w = 3$); ##431, . . . , 522 are of Type $W4$; ##530, . . . , 691 are of Type $W5$; ##701, . . . , 730 are of Type $W6$ (i.e., $w = 6$ *or* 7).

(b) The *total discharging value v* of the L-situation. By this we mean the sum of $d(E)$ (where E is the distinguished edge) plus 30 for every other edge from the pivot to a V_5 in the L-situation. In some cases (##530, 549, 550, 701, . . . , 728) the L-situation contains another L-situation with the same pivot but different distinguished edge, say F; in all of these cases, F is L5; then we add 50 for F instead of 30. We also distinguish v-values in increments of ten. For instance, within Type $W5$,

   #530   has $v = 120$;
##531, . . . , 541 have $v = 100$ or 95;
##549, 550   have $v = 90$;
##551, . . . , 553 have $v = 80$;
##561, . . . , 620 have $v = 70$ or 65;
   #621   has $v = 60$;
##622, . . . , 624 have $v = 50$;
##631, . . . , 691 have $v = 40$ or 35.

3. The set $\mathscr{U}$ of reducible configurations

If the discharging procedure $\mathscr{P}(\mathscr{T}, \mathscr{S}, \mathscr{L})$ is applied to a triangulation Δ (as described in Section 2) then it may or may not *completely discharge* Δ, i.e., it may or may not be true that $q(V) \leq 0$ for every vertex V of Δ. Now suppose that Δ^* is a triangulation which does not contain any configuration belonging

to the set $\mathcal{U}$ of 1834 configurations[4] presented in Part II of this paper, i.e., that Δ^* *avoids* $\mathcal{U}$. Then we shall prove the following.

DISCHARGING THEOREM FOR $\mathcal{P}(\mathcal{T}, \mathcal{S}, \mathcal{L})$, $\mathcal{U}$. *If* Δ^* *avoids* $\mathcal{U}$ *then* $\mathcal{P}(\mathcal{T}, \mathcal{S}, \mathcal{L})$ *completely discharges* Δ^*.

If the Euler characteristic χ of Δ is positive then, by (2.2), no discharging procedure can completely discharge Δ and we obtain the following.

COROLLARY. *If* $\chi > 0$ *then* Δ *cannot avoid* $\mathcal{U}$. *In particular, every planar triangulation* ($\chi = 2$) *contains at least one member of* $\mathcal{U}$.

Since every member of $\mathcal{U}$ is four color reducible (see Part II) in the sense that it cannot be contained in (i.e., immersed into) any minimal 5-chromatic planar triangulation, this implies that 5-chromatic planar triangulations do not exist and we have the main result.

FOUR COLOR THEOREM. *Every planar triangulation is (vertex-) colorable with four colors.*

Proof of the Discharging Theorem for $\mathcal{P}(\mathcal{T}, \mathcal{S}, \mathcal{L})$, $\mathcal{U}$. It follows immediately from the definition of $\mathcal{P}$ that $q(V_6) = 0$ for every 6-vertex V_6 of Δ^*. Thus it remains to be proved that

(A) $q(V_5) \le 0$ for every 5-vertex V_5 of Δ^* and
(B) $q(V_k) \le 0$ for every major vertex V_k of Δ^* ($k \ge 7$).

Proof of (A).   First, we prove some preliminary lemmas on T-dischargings.

LEMMA (5-6-6). *If a 5-vertex* V *of* Δ^* *has three consecutive neighbors of degrees* 5, 6, 6 *respectively, then a T-discharging of* 20 *leaves* V *across the* 6-6 *edge (see Figure* 6).

Proof.   Assume in an arbitrary triangulation Δ, V is a 5-vertex with consecutive neighbors of degrees 5, 6, 6 so that no T2-discharging leaves V (in Δ) across

---

Figure 6

the 6-6 edge, say E. Then, by definition of the T-discharging, none of the T2-situations $T2\#\#5$, 6, 7 is contained in Δ so as to induce a T-discharging from V across E. Thus Δ contains one of the four configurations of Figure 7 and hence one of the four members of $\mathcal{U}$ which are circled in Figure 7. Thus $\Delta \neq \Delta^*$ which proves the lemma. ■
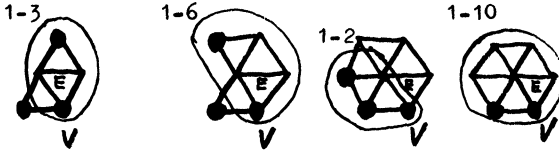
Figure 7

LEMMA (6-6-6).   *If a 5-vertex V of Δ^* has three consecutive 6-neighbors then a T-discharging of at least 10 leaves V across each of the two 6-6 edges (see Figure 8).*

Figure 8

LEMMA (5^5-7-6-6).   *If a 5-vertex V of Δ^* has four consecutive neighbors of degrees 5, 7, 6, 6 respectively so that another V_5 is adjacent to the 5- and 7-neighbors then a T-discharging of at least 10 leaves V across the 6-6 edge (see Figure 9).*

Figure 9

LEMMA (6-6).   *If a 5-vertex V of Δ^* is adjacent to a 6-6 edge so that no T-discharging leaves V across E then Δ^* contains the configuration of Figure* 10 (*with V and E identified to "V" and "E" as marked in Figure* 10).

Figure 10

The proofs of Lemmas (6-6-6), (5^5-7-6-6), and (6-6) are analogous to the proof of Lemma (5-6-6) (see the microfiche supplement).

LEMMA (S^+).   *If $f: S \to \Delta^*$ is an immersion of an S-situation S into Δ^* (which respects the degree specifications) then f can be extended to an immersion $f^+: S^+ \to \Delta^*$ of the enlarged S-situation S^+ (as drawn in Table* 1) *into Δ^* (so that f^+ also respects the degree specifications).*

This follows immediately by inspection of Table 1 (and the set $\mathcal{U}$ in Part II).

Now we consider the discharging procedure $\mathscr{P}(\mathscr{T}, \mathscr{S}, \emptyset)$ (which uses our T- and S-situations but no L-situations) and we denote the charge distribution which is obtained by this procedure by q_{TS}. Then we have the following.

$q_{TS}(V_5)$-LEMMA.   *If V is a 5-vertex of Δ^* so that $q_{TS}(V) > 0$ then one of the cases indicated in Table 3 applies, i.e., one of the configurations $CTS\#\#01, \ldots$, 33 drawn in Table 3 is contained in Δ^* with its central V_5 identified to V and so that S-situations are attached to the edges marked E, F, G, or H as indicated in Table* 3.

Proof.   This is proved by straightforward enumeration of all possible cases of $q_{TS}(V_5) > 0$ for $V_5 \in \Delta$ (Δ arbitrary as in Section 2). Then all cases which imply the presence of a configuration of $\mathcal{U}$ are deleted and the remaining cases are found in Table 3.

We denote by μ the number of major neighbors of V. Then we must consider the cases $\mu = 0, 1, 2, 3, 4, 5$. If $\mu = 0$, we have V surrounded by minor vertices which yields a member of $\mathcal{U}$ in every case. If $\mu = 1$, the only cases in which no member of $\mathcal{U}$ occurs are $CTS\#\#01, 02,$ and 03. The T2-arrows drawn in $CTS\#\#01$ and 03 indicate the T-dischargings according to Lemma (5-6-6). If $\mu = 2$ then at least one S-situation must be attached (since otherwise two R-dischargings would yield $q_{TS}(V_5) \le 0$); moreover, if only one S-situation is attached then the sum of its discharging value $d(E)$ and the values of the T-dischargings which may be implied by Lemmas (5-6-6), (6-6-6), or (5^5-7-6-6) must be smaller than 30 (since otherwise the remaining R-discharging would yield $q_{TS}(V) \le 0$). This leaves the cases $CTS\#\#04, \ldots, 13$. The case $\mu = 3$

$\mu = 1$

CTS #01

CTS #02

CTS #03

#002 or 003 at E

$\mu = 2$

CTS #04

#004, 007,
or 008
at E

CTS #05

#005 at E

CTS #06

one of
##011,...,016
at E

CTS #07

#017
at E

CTS #08

one of ##
021,...,
025
at E

CTS #09

one of ##051,...,054, 056,
061,...,068, 076, 079,...,083,
095, 102, 103, 117,...,119,
124, 146 at E

CTS #10

one of ##
171,
180, 181,
191,
199
at E

CTS #11

one of ##171,
174, 180, 181,
191, 192, 198,
199 at E

CTS #12

#151
at E

CTS #13

one of
174,
192, 198
at E

CTS #11.1

Table 3, page 1

$\mu = 3$

one of *) (013, 261n),
(011, 265r), (014, 261n),
(011, 266r), (015, 261n),
(011, 274), (016, 261n),
(011, 325r) at (E, F)

CTS #14

one of ##
032,...,035 at E

#248r
at F

CTS #15

one of
(052, 231n), (125, 231n), (118, 235n),
(052, 234n), (125, 234n), (118, 236n),
(052, 235n), (125, 235n),
(052, 236n), (125, 237n),
(052, 237n) at (E, F)

CTS #16

one pair out of
{052, 054, 055} × {231n, 234n,
     235n, 236n, 237n}
or {118} × {235n, 236n}
at (E, F)

CTS #17

#151 at E

#261n at F

CTS #18

#001 at E

one of
##011,
 · · · ,
016
at F

CTS #19

one pair out of

{171, 174, 180, 181, 191, 192,
198, 199} × {011,..., 016}
at (E, F)

CTS #20

#002 or 003 at E

#164 or
167 at F

CTS #21

#006 at E

one of
##011,
 · · · ,
016
at F

CTS #22

one pair of
{021, 022} ×
{002, 003}
or (030, 002)
at (E, F)

CTS #23

Table 3, page 2

*) The "r" in 265r, etc., means the reflected (mirror image) configuration
#265 from Table 1; "n" means non-reflected. See the definitions in Section 2.

one pair out of
$\{021, 022, 023,$
$025\}^2$
   at (E, F)

10 10
E 4|0 F
X

CTS #24

#151 at F;
one of ##171,...,173, 175,
176, 181,...,190, 193,
194, 199,...,204,
211,...,221, 224,...,229
   at E

20(25) 0
E 40 F
(35) X

CTS #25

#151 at F
one of ##
161,...,167
at E

10 0
E 5|0 F
X

CTS #26

#151
at E

#151
at F

0 0
E 6|0 F
X

CTS #27

one pair
out of
$\{161,...,$
$167\}^2,$
except (161,
161), at (E,F)

10 10
E 4|0 F
X

CTS #28

$\mu = 4$

one triplet out of
$\{011,...,016\} \times \{S2 \text{ (arbitrary of Type 5-major)}\}$
$\times \{231n, 234n,..., 237n\}$
   at (E, F, G)

0(5) F 0
E 40 G
(35) X

CTS #29

one triplet out of
$\{012, 015, 016\} \times \{011\} \times \{265r, 266r, 274, 325r\},$
or $\{011,...,016\} \times \{015, 016\} \times \{261n\}$
or $\{012,...,016\} \times \{013, 014\} \times \{261n\}$
      at (E, F, G)

0(5) F
0(9) 20
E 40 G
(35) X

CTS #30

one out of
$\{011,...,016\} \times$
$\{032,...,055\}$
$\times \{248r\}$
   at (E, F, G)

0 F
10
E 4|0 G
X

CTS #31

#151 at F

#261n at G

#151
at E

0 0 F
E 40 30 G
X

CTS #32

$\mu = 5$

one of 231r, 234r,...,237r at E,
one of 231n, 234n,...,237n at H,
any arbitrary S1 (of Type major-major) at F,
any arbitrary S1 (of Type major-major) at G

F 0 0 G
0 0
E 4|0 H
X

CTS #33

Table 3, page 3

yields the largest tree of sub-cases. In this case, at least two S-situations must be attached and at least one of them must be of Class S0 or S1. But if the sum of the discharging values $d(E) + d(F)$ is not smaller than 30 then a third S-situation must be attached. All those cases which do not yield configurations of $\mathcal{U}$ involve either an S0- or two S1-situations. (All combinatorial details of the enumeration of sub-cases are in the supplement to this paper.)

In the case $\mu = 4$, two S-situations must be attached at consecutive edges E, F so that $d(E) + d(F) < 30$ and so that E is of Type minor-major while F is of Type major-major. Then a third S-situation must be attached at an edge G, and if $d(E) + d(F) + d(G) \geq 30$ then a fourth S-situation is required. All cases which do not yield configurations of $\mathcal{U}$ involve either two S0- and one S2-, or one S0- and two S1-situations. Finally, if $\mu = 5$, all edges from V are of Type major-major. It is convenient to first consider all adjacent attachments of an S0 and an S0 or S1; only a few of them do not yield configurations of $\mathcal{U}$. Then one finds that all consecutive triplets $(S0, S0, S0)$, $(S0, S0, S1)$, and $(S1, S0, S1)$ yield configurations of $\mathcal{U}$. It is then easy to enumerate the remaining cases; the only ones which do not yield configuration of $\mathcal{U}$ involve four consecutive S-situations in the order $S0$, $S1$, $S1$, $S0$. ∎

The proof of (A) will be completed by the following.

L-LEMMA.   *Each of the configurations $CTS\#\#01, \ldots, 33$ of Table 3 (with S-situations attached as indicated in Table 3) contains an L-situation (of Table 2) which is attached to the edge marked X and the discharging value of which is large enough to yield $q(V) \leq 0$ for the central V_5.*

The proof of this is immediate by inspection. For the details see the supplement to Part I (as presented on microfiche cards in the back cover of this issue). In many cases the L-situation is identical to the configuration of Table 3 (with S-situations attached). ∎

Proof of (B).   Again we begin with some lemmas the proofs of which are obtained by straightforward case enumeration.

LEMMA (T).   *If V_k is a major vertex of Δ^* which is adjacent to a 6-6 edge E then V_k receives at most one T-discharging across E.*

LEMMA (T2, T2).   *If V_k is a major vertex of Δ^* which receives T 2-dischargings across two consecutive 6-6 edges E, F, then one of the two configurations of Figure 11 is contained in Δ^* (with V_k, E, F, as indicated in Figure 11).*

T2T2*n   T2T2*r
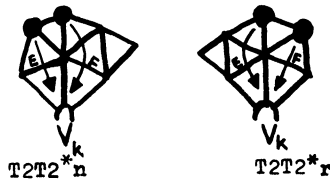
Figure 11

LEMMA (T2, T2, T2). *If V_k is a major vertex of Δ^* then V_k does not receive T2-dischargings across three consecutive 6-6 edges.*

LEMMA (T, T2, T2, T). *If a major vertex V_k of Δ^* receives T-dischargings across four consecutive 6-6 edges E, F, G, H then the dischargings across the second and third edges, F and G, are not both T2.*

LEMMA (60 OR 50, T). *If a major vertex V_k of Δ^* receives an L-discharging of 60 or 50 along an edge X and receives a T-discharging across a 6-6 edge E next to X (i.e., so that the 5-vertex of X and the 6-vertices of E are consecutive) then one of the two cases (a), (b) indicated in Figure 12 occurs.*

LEMMA (60 OR 50, T2, T2). *If a major vertex V_k of Δ^* receives an L-discharging of 60 or 50 and a T2-discharging across a 6-6 edge E next to X (see Figure 12a) then V_k cannot receive a second T2-discharging across a 6-6 edge F consecutive to E.*

(a) one of ##411, 441, 491,..., 495 at X   (b) one of ##411, 441, 492, 493, 494 at X

(c) 2 times #441   (d) #441 and one of ## 551, 552, 720

(f) #432 and #403   (g) #432 and #431 or 465   (h) 2 times #432
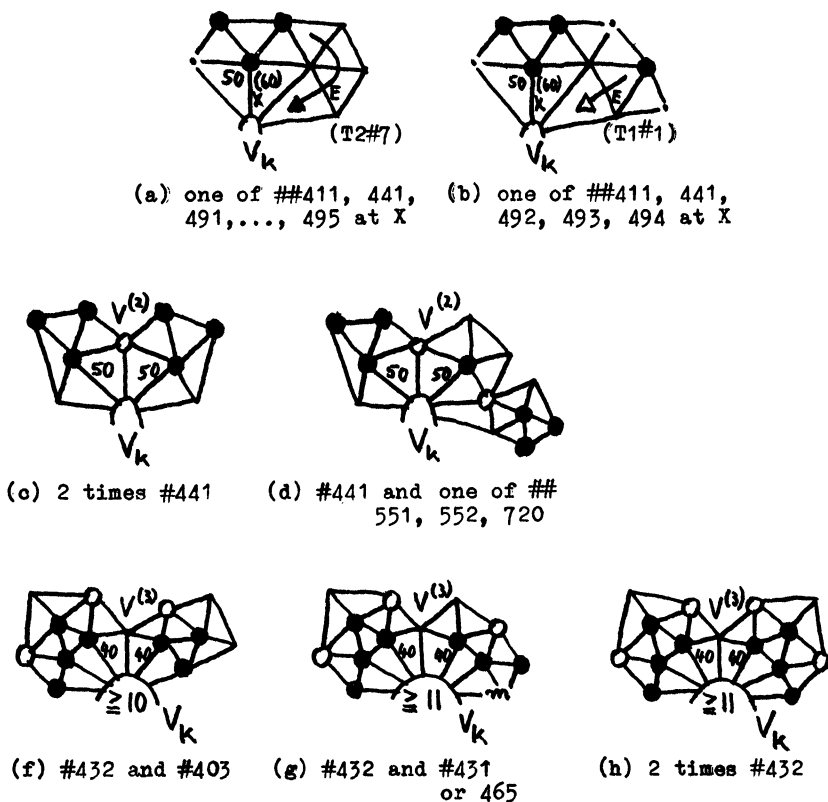
Figure 12

LEMMA (60 OR 50, ·, 60 OR 50). *If $V_5^{(1)}$, $V^{(2)}$, $V_5^{(3)}$ are three consecutive neighbors of a major vertex V_k of Δ^* such that V_k receives an L6- or L5-discharging from $V_5^{(1)}$ and another L6- or L5-discharging from $V_5^{(3)}$ then one of the two cases* (c), (d) *indicated in* Figure 12 *occurs* (*in particular, both dischargings are* L5 *and* $\deg(V^{(2)}) = 7$).

LEMMA (5, L, ·, L, 5). *If $V_5^{(1)}$, $V_5^{(2)}$, $V^{(3)}$, $V_5^{(4)}$, $V_5^{(5)}$ are five consecutive neighbors of a major vertex V_k of Δ^* such that V_k receives L-dischargings from $V_5^{(2)}$ and from $V_5^{(4)}$ (and such that $V_5^{(1)}$ and $V_5^{(5)}$ are of degree 5) then one of the three cases* (f), (g), (h) *indicated in Figure 12 occurs* (*in particular,* $\deg(V^{(3)}) = 6$ *and* $k \geq 10$).

LEMMA (5, L, 5). *If $V_5^{(1)}$, $V_5^{(2)}$, $V_5^{(3)}$ are three consecutive 5-neighbors of a major vertex V_k of Δ^* then no L-discharging can go from the second 5-vertex $V_5^{(2)}$ to V_k.*

LEMMA (5, L). *If $V_5^{(1)}$, $V_5^{(2)}$ are two adjacent 5-neighbors of a major vertex V_k of Δ^* then no L5- or L6-discharging can go from $V_5^{(1)}$ or $V_5^{(2)}$ to V_k (but L4-dischargings are possible).*

LEMMA (5, L, T2). *If $V_5^{(1)}$, $V_5^{(2)}$, $V_6^{(3)}$, $V_6^{(4)}$ are four consecutive neighbors of a major vertex V_k of Δ^* such that an L-discharging goes from $V_5^{(2)}$ to V_k and a T2-discharging goes to V_k across the edge E between $V_6^{(3)}$ and $V_6^{(4)}$ then the con-configuration of Figure 13 occurs* (*i.e.,* 40#432 *is attached and* $k \geq 10$).
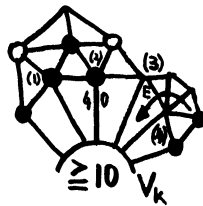
Figure 13

LEMMA (5, L, ·, 60 or 50). *If $V_5^{(1)}$, $V_5^{(2)}$, $V^{(3)}$, $V_5^{(4)}$ are four consecutive neighbors of a major vertex V_k of Δ^* such that V_k receives an L-discharging from $V_5^{(2)}$ and an L6- or L5-discharging from $V_5^{(4)}$ then the discharging from $V_5^{(4)}$ is induced by the L-situation* 50#441 (*i.e., is* L5) *so that* $\deg(V^{(3)}) = 7$.

LEMMA (50 OR 60, T, 50 OR 60). *If $V_6^{(1)}$, $V_6^{(2)}$, $V_6^{(3)}$, $V_6^{(4)}$ are four consecutive neighbors of a major vertex V_k of Δ^* such that from each of $V_5^{(1)}$, $V_5^{(4)}$ an L-discharging of value greater than 40 goes to V_k and such that a T-discharging goes to V_k across the edge E between $V_6^{(2)}$ and $V_6^{(3)}$ then* T1#1 *is attached at E* (*and thus the T-discharging transfers a charge of* 10 *to* V_k).

The above lemmas enable us to compute an upper bound for the sum of all

dischargings which go to a major vertex V_k of Δ^*; we denote this sum by $d(V_k)$ and the number of V_5-neighbors of V_k by $v(V_k)$.

UPPER BOUND LEMMA FOR $d(V_k)$. *Let V_k be a major vertex of Δ^*; let v be the number of neighbors of V_k of degree 5 and let k be the degree of V_k. Then the following inequalities hold.*

(3.1) $$d(V_k) \leq 30k - 7.5(k - v).$$

If the configuration of Figure 13 does NOT occur (with pivot identified to V_k) then

(3.2) $$d(V_k) \leq 30k - 10(k - v).$$

Proof. Let $V^{(1)}, \ldots, V^{(k)}$ be the neighbors of V_k in some clockwise cyclic order. We assign to each $V^{(i)}$ ($i = 1, \ldots, k$) a *contribution value* $c^{(i)}$ so that

$$c^{(i)} = \begin{cases} 30 & \text{if } V^{(i)} \text{ is of degree 5,} \\ c_*^{(i)} + c_{**}^{(i)} & \text{if } V^{(i)} \text{ is of degree greater than 5,} \end{cases}$$

where the $c_*^{(i)}$ and $c_{**}^{(i)}$ are defined as follows (with indices modulo k).

(a) If $V^{(j)}$ is a 5-vertex which L4-, L5-, or L6-discharges to V_k and if neither of $V^{(j-1)}$, $V^{(j+1)}$ is a 5-vertex then $c_{**}^{(j-1)} = c_*^{(j+1)} = 5$, 10, or 15, respectively, except in the three cases of Figure 14 where $c_{**}^{(j-1)}$ and $c_*^{(j+1)}$ are defined as indicated below the drawings.

$$c_{**}^{(j-1)} = 0, \quad c_*^{(j+1)} = 20 \qquad c_{**}^{(j-1)} = 0, \quad c_*^{(j+1)} = 20 \qquad c_{**}^{(j-1)} = 15, \quad c_*^{(j+1)} = 5$$
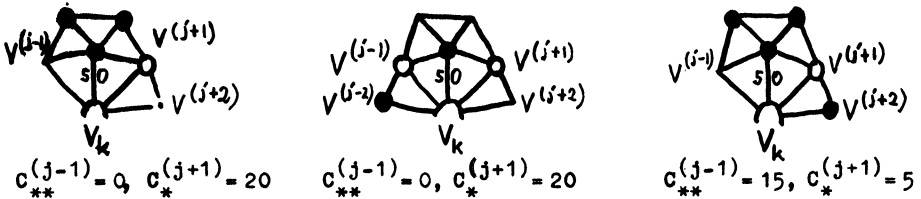
(analogous definitions for the mirror images of these configurations)

Figure 14

(b) If $V^{(j)}$ and $V^{(j+1)}$ are 5-vertices such that $V^{(j)}$ L4-discharges to V_k then $c_{**}^{(j-1)} = 10$. Correspondingly, if $V^{(j-1)}$ and $V^{(j)}$ are V_5's and $V^{(j)}$ L4-discharges to V_k then $c_*^{(j+1)} = 10$.

(c) If $V^{(j)}$ and $V^{(j+1)}$ are 6-vertices such that a T1-discharging goes to V_k across the edge E between $V^{(j)}$ and $V^{(j+1)}$ then $c_{**}^{(j)} = c_*^{(j+1)} = 5$.

(d) If $V^{(j)}$ and $V^{(j+1)}$ are 6-vertices such that a T2-discharging goes to V_k across the edge E between $V^{(j)}$ and $V^{(j+1)}$ then $c_{**}^{(j)} = c_*^{(j+1)} = 10$, except in the three cases of Figure 15, where $c_{**}^{(j)}$ and $c_*^{(j+1)}$ are defined as indicated below the drawings.

$$c^{(j)}_{**} = 15, \quad c^{(j+1)}_{*} = 5 \qquad c^{(j)}_{**} = 15, \quad c^{(j+1)}_{*} = 5 \quad c^{(j)}_{**} = 12.5, \quad c^{(j+1)}_{*} = 7.5$$

(analogous definitions for the mirror images of these configurations)
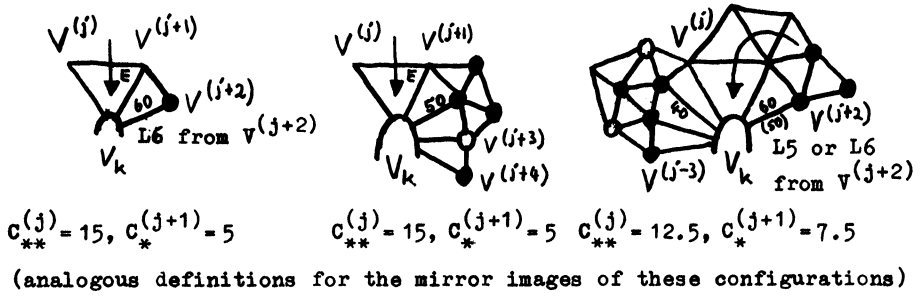
Figure 15

(e) If for some i, $c^{(i)}_{*}$ or $c^{(i)}_{**}$ is not defined by (a), (b), (c), or (d) then it is defined to be zero.

Now, by the above lemmas we have

(3.3) $$\sum_{i=1}^{k} c^{(i)} \geq d(V_k);$$

(3.4) $$c^{(i)} \leq \begin{cases} 22.5 & \text{if } V^{(i)} \text{ is not a } V_5 \\ 20 & \text{if } V^{(i)} \text{ is neither a } V_5 \text{ nor a } V_6 \text{ in the configuration of Figure 13.} \end{cases}$$

This immediately implies (3.1) and (3.2). ∎

COROLLARY. (Notation as in the proof of the Upper Bound Lemma.) *If the configuration of Figure* 13 *does not occur and if there is some index l such that* $c^{(l)} < 20$ *then*

(3.5) $$d(V_k) \leq 30k - 10(k - v) - 10.$$

Proof. This follows from (3.3) and (3.4) since $\sum_{i=1}^{k} c^{(i)}$ is an integral multiple of 10.

Now it is easy to prove (B) for all vertices V_k of degree $k \geq 11$, using the fact that

(3.6) $$q(V_k) = d(V_k) - 60(k - 6).$$

If $k \geq 12$ then (3.1) and (3.6) imply immediately that $q(V_k) \leq 0$. If $k = 11$ then (3.1) and (3.6) imply that if $q(V_{11}) > 0$ then $v \geq 8$; but if $v = 8$ then, by (3.2), the configuration of Figure 13 must occur. On the other hand, $v \geq 10$ is ruled out in Δ^* because of 15-34 in $\mathcal{U}$. Now it is easy to check that in all remaining cases $q(V_{11}) \leq 0$ or 15-34 or 15-35 occur.

It remains to prove (B) for $k = 7, 8, 9$, and 10.

We consider the discharging procedure $\mathcal{P}(\mathcal{T}, \emptyset, \mathcal{L})$ (i.e., we ignore the S-situations) and we denote the corresponding charge function by q_{TL}. Then we have the following.

$q_{TL}(V_k)$-LEMMA. *If V_k is a major vertex in Δ^* of degree $k \geq 7$ such that $q_{TL}(V_k) > 0$ then one of the cases $CTL\#\#1, \ldots, 152$ of Table 4 applies, in the sense that the central major vertex is identified to V_k and that L-situations are attached as indicated to those edges which are marked by large discharging values but to no other edges incident to V_k (the small discharging values which are written at some edges are to be ignored in this context).*

The proof of this lemma is by straightforward case enumeration for $k = 7, 8,$ 9, and 10 (see the microfiche supplement).

For $k = 7$ one must consider the following neighborhoods of a V_7.

(7.0) No V_5-neighbor but T-dischargings of total value greater than 60.

(7.1) One V_5-neighbor which is R-, L4-, L5-, or L6-discharging to the V_7 and T-dischargings of total value greater than 30, 20, 10, or 0, respectively.

(7.2) Two V_5-neighbors and a T- or L-discharging.

(7.3) Three or more V_5-neighbors to the V_7.

Those cases in which no member of $\mathcal{U}$ occurs are listed in Table 4. We remark that all cases in which an L-situation of width $w \geq 4$ occurs (i.e., all but the first twelve situations in Table 2) yield likely to be reducible configurations if the pivot is a V_7; but some of these configurations are of ring size greater than 14. For this reason we preferred to include the cases $CTL\#\#76, 77$ in Table 4 rather than try to reduce the corresponding 15-ring configurations in order to include them in $\mathcal{U}$.

For $k = 8$, which is the case of greatest combinatorial complexity, we must consider the following V_8-neighborhoods.

(8.1, 2, 3) One, two, or three V_5-neighbors and correspondingly strong T- and/or L-dischargings. The case (8.0), i.e., T-dischargings of total value greater than 120, can be ruled out by Lemmas (T), (T2, T2, T2), and (T, T2, T2, T).

(8.4) Four V_5-neighbors and a T- or L-discharging.

(8.5) Five or more V_5-neighbors.

The only cases which involve an L-situation of width $w \geq 5$ and do not yield a member of $\mathcal{U}$ are $CTL\#138$ and $\#139$ in Table 4.

For $k = 9$ and 10, the cases of $q_{TL}(V_k) > 0$ turn out to require so relatively many V_5-neighbors to the V_k that the total number of cases to be considered is smaller than for $k = 8$.

S-LEMMA. *In all cases $CTL\#\#1, \ldots, 152$ of Table 4 there occur S-situations (shown by drawings in Table 1) which induce the small discharging values which are indicated in Table 4 so that in every case $q(V_k) \leq 0$ for the central V_k. (In some cases the S-situations may induce even smaller discharging values than indicated in Table 4.)*

The proof of the S-Lemma is immediate by inspection. This completes the proof of (B) and thus of the Discharging Theorem.

K. APPEL AND W. HAKEN

CTL #1 CTL #2 CTL #3 CTL #4 CTL #5
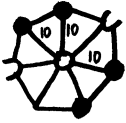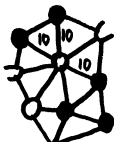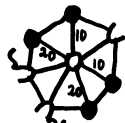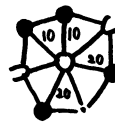
CTL #6 CTL #7 CTL #8 CTL #9 CTL #10
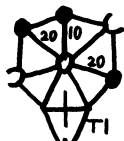
CTL #11 CTL #12 CTL #13 CTL #14 CTL #15
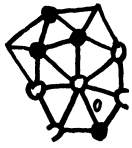
CTL #16 CTL #17 CTL #18 CTL #19 CTL #20
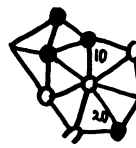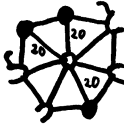
CTL #21 CTL #22 CTL #23 CTL #24 CTL #25

Table 4, page 1
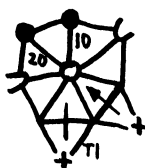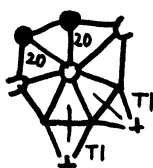
CTL #26   CTL #27   CTL #28   CTL #29   CTL #30
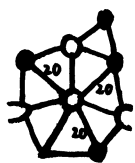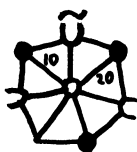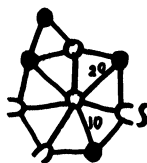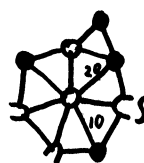
CTL #31   CTL #32   CTL #33   CTL #34   CTL #35
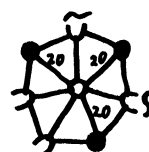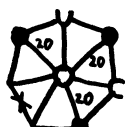
CTL #36   CTL #37   CTL #38   CTL #39   CTL #40
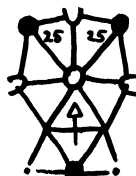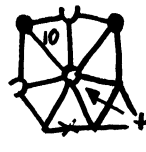
CTL #41   CTL #43   CTL #44   CTL #45
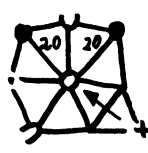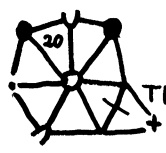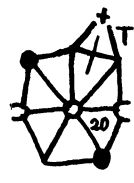
CTL #46   CTL #47   CTL #48   CTL #49   CTL #50

Table 4, page 2

CTL #51   CTL #52   CTL #53   CTL #54   CTL #55

(#403)

#401 or
402 at Y

#401, 402, or 403 at X

CTL #56   CTL #57   CTL #58   CTL #59   CTL #60

#401, 402, or 403 at X

#401 or 402 at Y

CTL #61   CTL #62   CTL #63   CTL #64

(#411)

(#403)

CTL #65   CTL #66   CTL #69   CTL #70

#421,...,425,
or 427 at Z

(#428)

#421,...,425, 427, or 428
at U

CTL #71   CTL #72   CTL #74   CTL #75
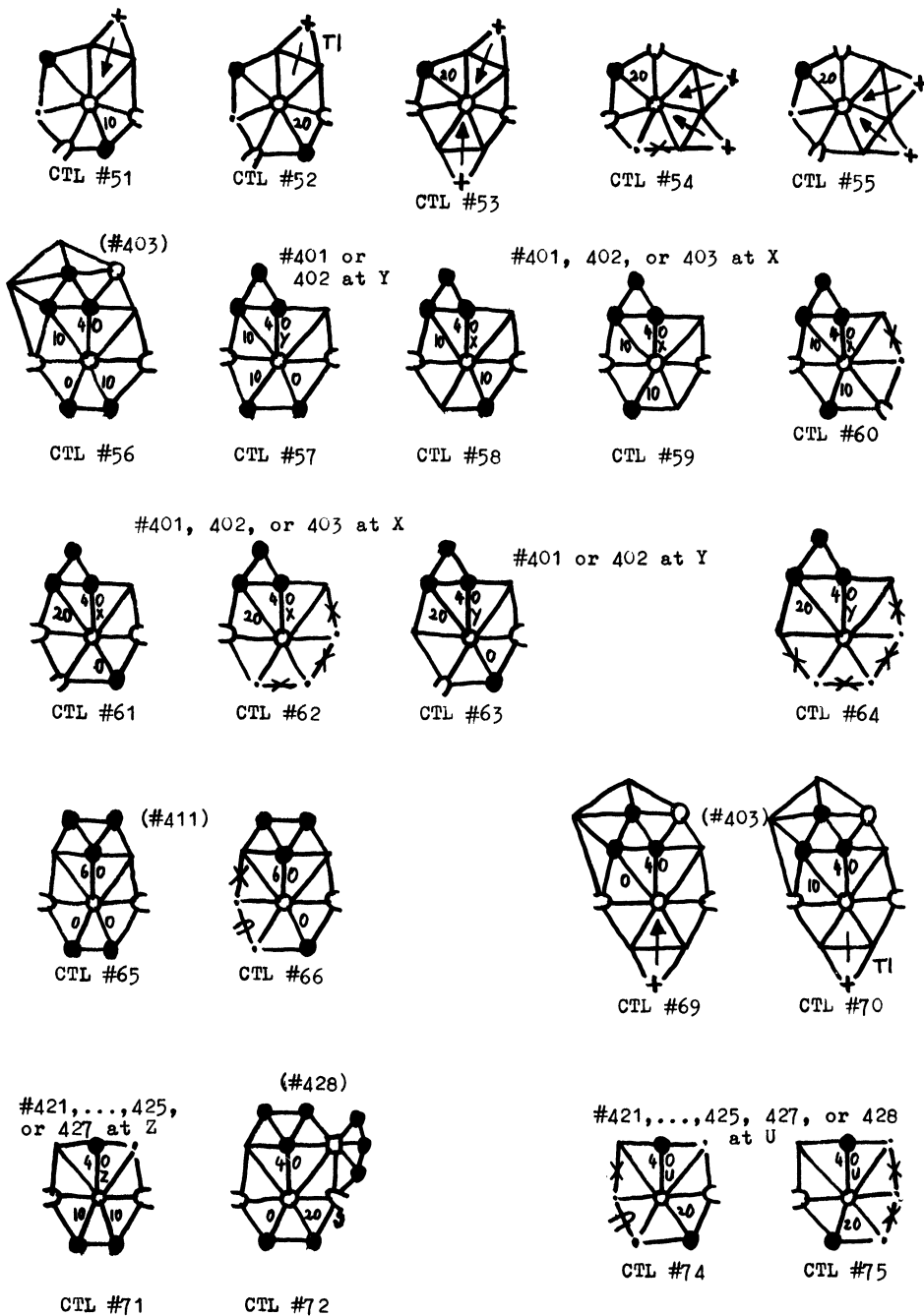
Table 4, page 3

one of ## 604,...,607
615, 616, 617 at Z   615, 616, or 617 at U
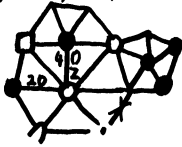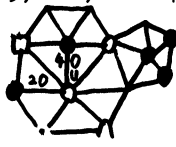
CTL #76   CTL #77
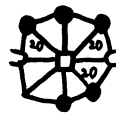
CTL #81   CTL #82   CTL #83   CTL #84   CTL #85

CTL #86   CTL #87
                             CTL #88   CTL #89   CTL #90

CTL #91   CTL #92   CTL #93   CTL #94

#401, 402, or   #401, 402, or
  403 at X   #401 or 402 at Y   (#403)   403 at X
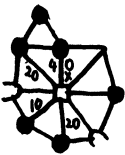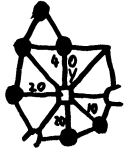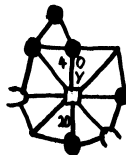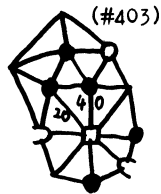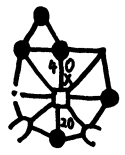
CTL #96   CTL #97   CTL #98   CTL #99   CTL #100

Table 4, page 4

#401, 402, or 403 at X
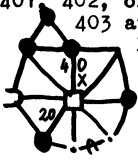
CTL #101

#401 or 402 at Y

CTL #102
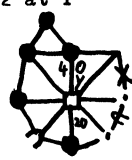
CTL #103

#401, 402, or 403 at X
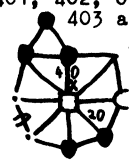
CTL #104

#402 or 403 at Z
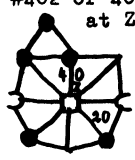
CTL #105

#401, 402, or 403 at X
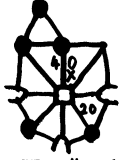
CTL #106
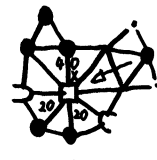
CTL #107
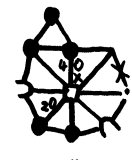
CTL #108

CTL #109

CTL #111

CTL #112
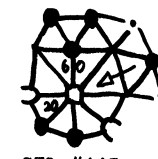
CTL #113

#421, 422, or 423 at U
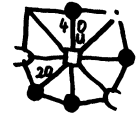
CTL #116

one of ## 421,...,428 at Q
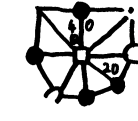
CTL #117

(#426)
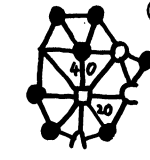
CTL #118
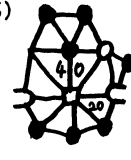
CTL #119

#401, 402, or 403 at X, #401, 402, or 403 at P

CTL #121

CTL #122

(2× #401)
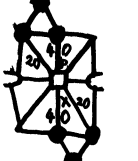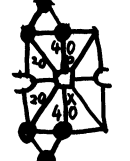
CTL #123

(2× #402)
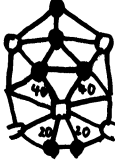
CTL #124

(#401 and #403)

CTL #125

Table 4, page 5

(#431)   (#432)

#431 or 432 at X

one of ##
433,...,437 at Y

CTL #126   CTL #127   CTL #128   CTL #129   CTL #130

one of ##433,...,437 at Y   (#441)   one of ##
451,...,464 at
Z   (#465)

CTL #131   CTL #132   CTL #133   CTL #134   CTL #135

one of ##
466,...,481 at
U   (#475)   (#532)   one of ##580,
591, 592, 595,
611,...,613,
615,...,617 at Q

CTL #136   CTL #137   CTL #138   CTL #139

#401   or   402   at   P   (#403)

CTL #141   CTL #142   CTL #143   CTL #144   CTL #145

#431 or 432 at X   one of ##433,
...,436 at N   one of ##433,
...,437 at Y   (#437)

CTL #146   CTL #147   CTL #148   CTL #149   CTL #150

(#437)   Table 4, page 6

CTL #151   CTL #152

4. Probabilistic considerations

At first glance it may appear a very strange accident that Kempe's attempted proof of the Four Color Conjecture can be repaired by a method of such considerable combinatorial complexity whereas a moderately simple repair seems unlikely.

In this section we will present an argument based on elementary (and rather crude) use of probability which leads to the belief that it is overwhelmingly likely that there "must" exist an unavoidable set of reducible configurations with ring sizes n not exceeding 17, and further that it is very likely that there exists an unavoidable set with ring sizes $n \leq 14$, while it is unlikely that such a set with ring sizes $n \leq 12$ exists. Only for sets with $n \leq 13$ can no prediction be made.

The example of a map constructed by Edward F. Moore in March 1977 settles the existence question in the negative for $n \leq 11$ while the present paper settles it affirmatively for $n \leq 14$, so that only $n \leq 12$ and $n \leq 13$ remain open. While this paper may be a surprise for many mathematicians, taken together with the work of Moore, all it actually shows is that there are no surprises. Thus, for the time being, the probabilistic discussion we will now outline appears a sufficient explanation for the great, but limited, degree of difficulty of the Four Color Problem.

(a) *Likelihood of reducibility.* In what follows we shall consider D-reducibility (in the sense of Heesch [16]) as our paradigm for reducibility. The reader unfamiliar with the technical aspects of D-reducibility will find them in [16] and [27]. While C-reducibility is somewhat stronger, it does not appear sufficiently stronger to change the result by a difference of 1 in the required n and its parameters seem much more difficult to define. Thus, in this sub-section, reducible will mean D-reducible.

Given a configuration C of ring-size n, say for instance, $n = 13$, one may ask "How likely is C to be reducible?". We consider the ring R of thirteen vertices surrounding C and the 66,430 different classes of colorations of R. A certain percentage of these colorations, say x, can be extended through C; we call these colorations "initially good" or "good$_0$". We consider an arbitrary coloration c of R which is not good$_0$ and we ask "How good are our chances to convert this coloration into a good one by a simple Kempe-chain argument?", i.e., how likely is c to be good$_1$ (of chromatic distance one from the set of good$_0$ colorations)?

Consider a partition π of the four colors we permit into two pairs. The ring will be split into 2-colored components with respect to the pairs, and the number of components corresponding to the two pairs will be the same (since they occur in cyclically alternating order on R). This number may be 1, 2, 3, 4, 5, or 6; but for at least one of the three possible partitions the number of components corresponding to each pair must be either 5 or 6 and for at least one of the remaining two partitions it must be at least 4. We will restrict ourselves to partitions which yield 4, 5, or 6 pairs of two-colored components of R. Then

we have 7, 15, or 31, respectively, choices of Kempe interchanges by which we may derive a different coloration c' of R from c. When we consider the possible colorations of the exterior of R which induce the coloration c of R, we note that there are 14, 42, or 132 possible Kempe chain dispositions according as our partition π had 4, 5, or 6 pairs of components on R. If for *every* Kempe chain disposition, *some* Kempe interchange yields a $good_0$ coloration c' then c is $good_1$. For lack of better usable information we assume that the $good_0$ colorations are randomly distributed among all the colorations of R and comprise $100x$ percent of them. Then the probability that for each Kempe chain disposition some Kempe interchange is good is

$$y_4 = [1 - (1 - x)^7]^{14},$$

(4.1) $$y_5 = [1 - (1 - x)^{15}]^{42},$$

$$y_6 = [1 - (1 - x)^{31}]^{132},$$

respectively. These functions are plotted in Figure 16.

The graphs suggest that for $x \leq 10$ percent there is practically no chance for the Kempe chain argument to succeed; for $x = 20$ percent there is a very good
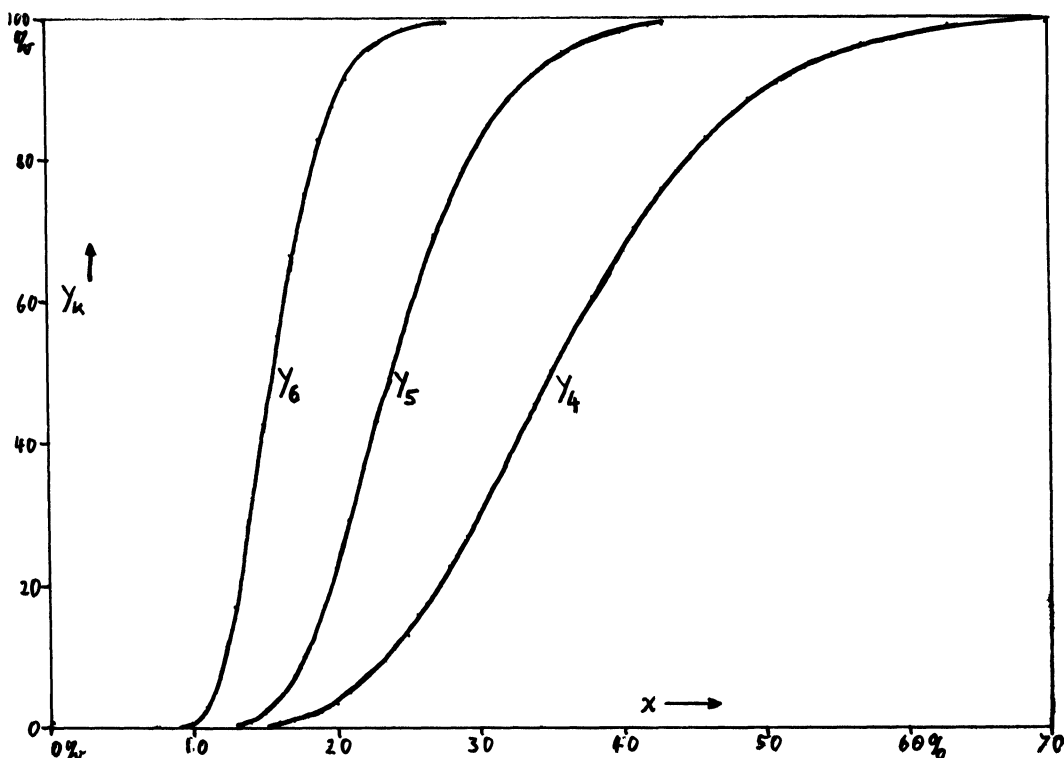
Figure 16: Probability of conversion vs. percentage of $good_0$ colorations

chance of success; and for $x \geq 30$ percent one should expect reducibility to be almost certain. For examples see the configurations (b), (c), and (d) in Figure 16* where g means the number of $good_0$ coloration classes of the ring R around the configuration. The configurations (b) and (c) with x-values of 11.4 and 12.1 percent are not D-reducible (but C-reducible) while (d) with an x-value of 14.6 percent is D-reducible.

Next we estimate the likelihood of reducibility as a function of the number m of vertices of the configuration C (not counting the vertices on the ring R). While it is not so easy to estimate what value of x is to be expected for given m and n, it is very plausible that for a fixed value of n, the average of the values of x will increase quite rapidly with m (since the total number of possible

degree-raising

(a)   27-28
   n = 12
   m = 9
   g = 3703
   x = 0.167
   D-reducible

(b)   43-23
   n = 13
   m = 9
   g = 7575
   x = 0.114
   not D-, but C-reducible

1-extension

(c)   52-22
   n = 13
   m = 9
   g = 8044
   x = 0.121
   not D-, but C-reducible

(d)
   n = 13
   m = 10
   g = 9680
   x = 0.146
   D-reducible
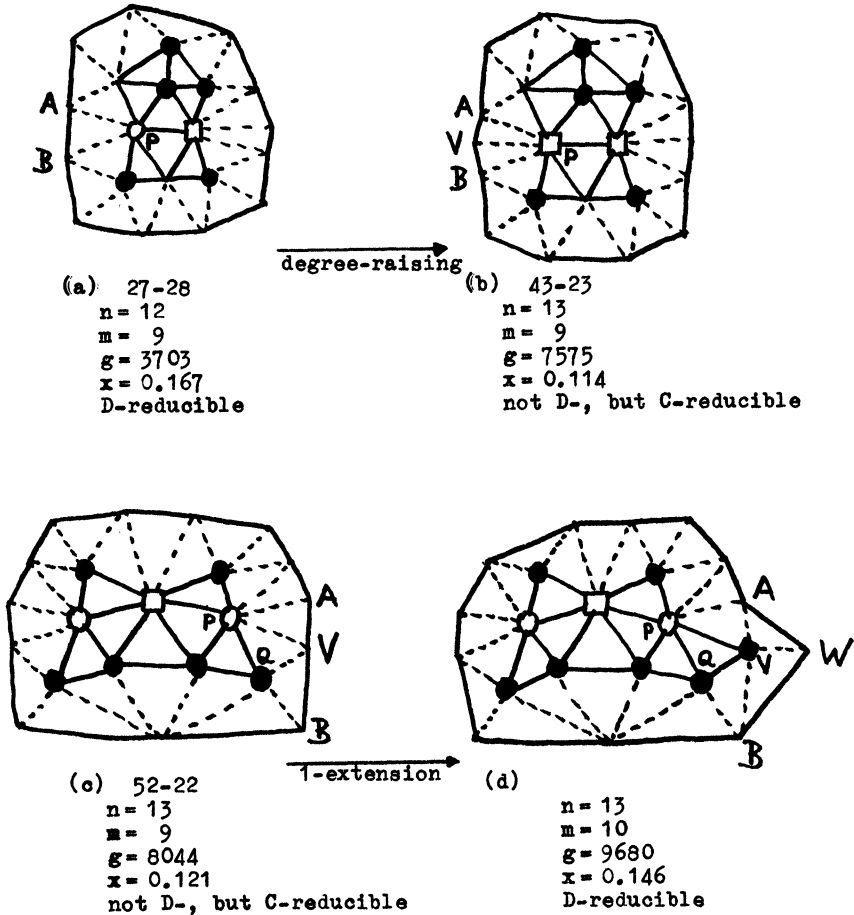
Figure 16*

colorations of C will increase considerably with m).[5] On this basis one might

***

[5] A reasonable estimate of the dependence of x on the parameters n and m can be obtained as follows. In order to estimate the dependence on n we raise the degree of a vertex P of a configuration C by one. Denote the configuration so obtained by C_{n+1}; for example let C and C_{n+1} be (a) and (b) in Figure 16*. We assume that there is a triangle ABP such that A and B are on the ring R around C (i.e., that P is at least a 2-legger vertex of C). In order to obtain C_{n+1} and the ring R_{n+1} around it we subdivide the edge AB of R by a new vertex V and correspondingly we subdivide the triangle ABP by a new edge from V to P (a new leg at P). Now we consider all the g_{n+1} good$_0$ colorations of R_{n+1}, each of them being extended over C_{n+1}. Pretending that we do not know any details about C_{n+1} we may estimate that 50 percent of these colorations will have the same color at A and at B. These colorations do not correspond to colorations of R, while the other 50 percent (where A has a color different from B) correspond to one-to-one to the g good$_0$ colorations of R. Thus we obtain the estimates

$$g_{n+1} = 2g \quad \text{and} \quad x_{n+1} = 0.66x.$$

(Note that the total number of colorations of R_{n+1} is almost precisely three times as large as the number of colorations of R.)

In order to estimate the dependence of x on m we add a ring-vertex V to a configuration C and give a degree specification to V so as to obtain a configuration, denoted by C_{m+1}, with the same n-value as C (this operation is called a 1-extension in [26]); for example let C and C_{m+1} be (c) and (d) in Figure 16*. We assume that the new vertex V has precisely two neighbors, say P and Q, in C and thus has degree-specification 5 in C_{m+1}. In R_{m+1} a new vertex, say W, occurs as a replacement for V in R. Now we consider all the g good$_0$ colorations of R, each one extended over C. Let the vertices A, P, Q, B lie in that order around V. Then we may estimate that in 50 percent of the colorations the vertex Q has a color different from A and that in 50 percent of those colorations the vertex B has the same color as A. Thus in 25 percent of all good$_0$ colorations of R, A and B have the same color. To each of these colorations correspond precisely two good$_0$ colorations of R_{m+1} (two choices of color for W in each case), while those 75 percent of the colorations with A and B colored differently correspond one-to-one to colorations of R_{m+1}. Thus it appears that we can expect g_{m+1} to be 25 percent larger than g. However if c is a good$_0$ coloration of R which gives A and B the same color then we have a likelihood of $2x$ percent that there is another good$_0$ coloration, say c', of R which agrees with c on all vertices of R except on V. In this case we would have only three different colorations of R_{m+1} corresponding to the two colorations c and c'. Thus we may estimate that g_{m+1} will be only $25(1 - x)$ percent larger than g.

Now if we consider two configurations C and C', can we expect that it is possible to change C into C' by a sequence of the above operations (degree-raisings and 1-extensions) and their inverses? If we assume that C and C' are not articulated then the answer will be "yes" in the majority of cases of interest to us. However, occasionally we may need a 1-extension where the vertex V is adjacent to precisely three vertices of C and correspondingly has degree specification 6 in C_{m+1}. In this case the increase in g-value can be estimated to be $37.5(1 - x)$ percent.

For configurations with n- and m-values of the order of magnitude in which we are interested we may now estimate that an increase in n by one yields a decrease in x by 33 percent and that an increase in m by one yields an increase in x by 23 percent. This explains the empirical observation that (for obstacle-free configurations) the likelihood of reducibility depends essentially on the difference $n - m$. If both n and m are increased by one then x will decrease by about 18 percent. This means for instance that an 11-ring configuration with 8 vertices will have an x-value of about 1.5 times the x-value of a 13-ring configuration with 10 vertices. But for $n = 11$ the curves y_5 and y_4 of Figure 16 play the same role which y_6 and y_5 play for $n = 13$. Therefore the likelihood of reducibility can be assumed to be about the same for both configurations, probably somewhat larger for the 13-ring than for the 11-ring configuration.

predict that there will be a "critical" value m' such that $m > m'$ means "likely to be reducible" while $m < m'$ means "likely to be irreducible."

Of course the above predictions should make use of the knowledge we do have and should be applied only to configurations which do not contain any of the known reduction obstacles (see [26; Chapters II, III]) and thus in particular are geographically good and without hanging pairs (cf. Section 1 of this paper). For small ring-sizes, $n = 6, \ldots, 11$, these configurations have been studied quite exhaustively (F. Bernhart [8], Allaire and Swart [2], Koch [20]). The results suggest that the "critical m-value" satisfies

$$(4.2) \qquad\qquad m' = n - 5.$$

For $n = 6, 7, 8, 9$, all configurations without reduction obstacles satisfy $m > n - 5$ and all of them are C- or D-reducible. (Note that the 5-5-5 triangle contains hanging pairs and is meant to be excluded.) For $n = 10$ the only such configuration with $m = n - 5 = 5$ (7-5665) is irreducible[6] and those of greater m are reducible. (The careful reader will note that a few of these are C- but not D-reducible but the vast majority are D-reducible. The very fact that C-reducibility does not seem to change the "critical value" argues for the reasonableness of working with D-reducibility.) For $n = 11$ there are 6 configurations with $m = n - 5 = 6$ and 4 of them are reducible; only one configuration with $m > n - 5$ could not be reduced (8-556655, $m = n - 4 = 7$). On this basis it appears reasonable to accept (4.2) for $n = 12, 13, 14, \ldots$ also; one might actually expect that for some higher n-values $m' = n - 6$ would be more appropriate. The computer results obtained so far (see [1], [9], [17], and Part II of this paper) do not indicate any reason to doubt (4.2).

In [4] the authors introduced a function

$$(4.3) \qquad\qquad \phi(C) = n - m - 3$$

and proved that an *arbitrary* configuration C with $\phi(C) \leq 0$ always contains a geographically good subconfiguration, say C^*, again with $\phi(C^*) \leq 0$. (The proof is easy by induction on m.) But unfortunately, the sub-configuration C^* may still contain a hanging pair and thus may be not "likely to be reducible." However, we have the following.

m-LEMMA. *Let C be an arbitrary configuration of ring size n containing m vertices (inside the ring.) Assume that*

$$(4.4) \qquad m > 3n/2 - 6, \quad \text{or equivalently}, \quad \phi < 3 - n/2.$$

Then C contains a subconfiguration C^* which again satisfies (4.4) such that C^* is geographically good without hanging pairs. Moreover, if A is an articulation

---

[6] Here "irreducible" means C-irreducible. E. R. Swart has found recently that a stronger reduction method due to Arthur Bernhart can be used to show that the configuration 7-5665 is reducible. It appears however that the reduction obstacles mentioned here remain valid also with respect to the improved reduction method.

vertex of C^* and if W_1 and W_2 are the two "wings" of A (i.e., the two configurations obtained from C^* by deleting A) then we have

$$m(W_i) \geq 3n(W_i)/2 - 6 \quad \textit{for } i = 1, 2.$$

Proof. It is easy to check that m cannot be 1, 2, or 3 (since we have excluded vertices of degrees smaller than five) and that the only possible C with $m = 4$ is Birkhoff's reducible 5-555 double-triangle; thus $C^* = C$. Now assume by induction that the lemma holds for all configurations of up to $m - 1$ vertices. Then we claim that the lemma holds also for C. If C has all the properties required for C^* then we let C^* be C and the proof is finished. If C contains a 4- (or more-) legger vertex V then deleting V yields a configuration C' with $m(C') = m - 1$ and $n(C') \leq n - 1$; thus C' satisfies (4.4) and by the inductive hypothesis the required sub-configuration C^* is found in C'.

If C contains a "bad" articulation vertex B then we claim that at least one of the wings $C_1, \ldots, C_k$ $(k \geq 2)$ at B, say C_1, satisfies (4.4). Then by the inductive hypothesis, C_1 contains the required C^*.

In order to prove the claim we assume the contrary. Thus

(4.5) $$m(C_i) \leq 3n(C_i)/2 - 6 \quad \text{for each } i = 1, 2, \ldots, k.$$

Moreover, if $k = 2$ and B is a 2-legger vertex, then for at least one of the two wings, say for C_1, we have $m(C_1) < 3n(C_1)/2 - 6$. Now we have

(4.6) $$m = \sum_{i=1}^{k} m(C_i) + 1$$

(4.7) $$n \geq \sum_{i=1}^{k} [n(C_i) - 2]$$

where the equality sign in (4.7) holds if and only if B is a k-legger vertex (i.e., precisely one "leg" lies between each two consecutive wings). Combining (4.6) with (4.5) and (4.7) we have

$$m \leq \frac{3}{2} \sum_{i=1}^{k} [n(C_i) - 4] + 1 \leq 3n/2 - 3k + 1$$

where in the case $k = 2$ not both equality signs can hold. But this contradicts the hypothesis (4.4); this finishes the proof. ∎

By the above lemma, a configuration C which satisfies (4.4) does not contain any known reduction obstacles and must be regarded as extremely likely to be reducible. (Note that it far surpasses our critical condition given earlier.) One may conjecture that *every* configuration which satisfies (4.4) is D-reducible; but we do not expect that this conjecture can be proved.

(b) *Likelihood of unavoidability.* For any given integer $n_0 \geq 5$, we may try to find integers r and ϕ_0 such that every triangulation (planar with vertex degrees ≥ 5) contains at least one configuration with $n \leq n_0$ and $\phi \leq \phi_0$ which is contained in the rth neighborhood of a vertex. This question has been stimulated by the work of Stromquist who proved in [26; Chapter IV] that every triangulation contains at least one configuration with $\phi \leq -1$ which is contained in the second neighborhood of a vertex and in which no vertex has degree greater than 30 (thus proving the existence of a finite, unavoidable set of geographically good configurations).

We consider the following "size classes of neighborhoods" in triangulations.

Size Class	Description
1	single vertex
2	edge (pair of adjacent vertices)
3	triangle (vertex with two consecutive neighbors)
4	double triangle (vertex with three consecutive neighbors)
5	triple triangle (vertex with four consecutive neighbors)
6	first neighborhood of a vertex (vertex with all neighbors)
7	neighborhood N_6 of Class 6 plus one triangle (with base in N_6)
8	first neighborhood of an edge
9	first neighborhood of a triangle
. . .	. . .
s	first neighborhood of a neighborhood N_{s-6} of Size Class $s - 6$ (for $s \geq 8$)
. . .	. . .

Each of these size classes has an "average n" and an "average m" (the average taken over all configurations of the size class in all triangulations). Since the "average vertex degree" is certainly six we may get an approximate idea of what these averages are by considering configurations (of the respective size classes) which consist of 6-vertices only. The corresponding ϕ-values are plotted in Figure 17 versus the n-values and are marked X; the size class numbers are written above the marks. From size class 15 on ($n \geq 21$) the ϕ-values lie below the line $\phi = 3 - n/2$, and thus the average configuration of this size class will almost certainly be D-reducible (see the m-Lemma).

Certainly, every triangulation will contain some configurations of each size class the ϕ-values of which are below the average; for by (1.2), every triangulation must contain vertices of degree 5 (i.e., with degrees substantially below the average of 6). For an estimate of these "unavoidable ϕ-values" we have considered configurations of the different size classes such that one vertex, as close to the center as possible, is a 5-vertex while all other vertices are of degree 6. The results are marked • in Figure 17. From $n = 17$ on the values are in the region of extremely high likelihood of D-reducibility.
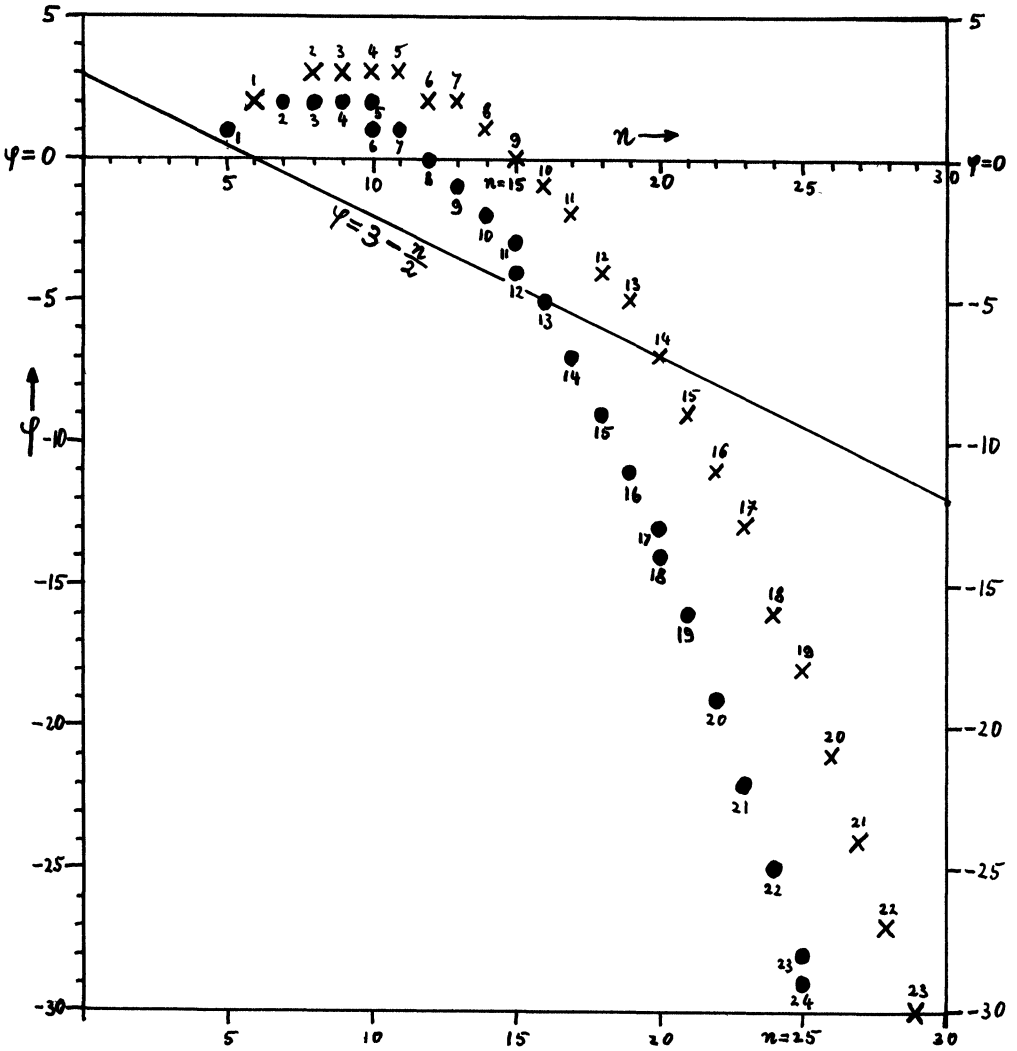
Figure 17:   Approximate values of average φ (marked ✗) and
unavoidable φ (marked ●) versus n;
size class as parameter

   On the other hand we must expect that there exist triangulations in which all
configurations with $n \leq 12$ have ϕ-values ≥ 0. It is then plausible to expect
further that in some of these triangulations, no configuration with $n \leq 12$ is
reducible. For instance, all these configurations might be such that iterated
removal of 4-legger vertices eventually yields a single V_5 ($\phi = 1$) or a 5-5-5
triangle ($\phi = 0$).

Similarly, we must expect the existence of triangulations in which every configuration with $n \leq 13$ has $\phi \geq -1$. But then it is not so easy to imagine that all these configurations are irreducible. Of course, after removal of hanging 5-5 pairs and of 4-legger vertices, a configuration with $\phi = -1$ and $n = 13$ may yield nothing better than a 5-5-5 triangle and be irreducible; but whether there may exist a triangulation such that this occurs in all cases seems to be not reasonably predictable.

The corresponding reasoning for $n \leq 14$, $n \leq 15$, and $n \leq 16$ tends more and more to the belief that every triangulation will contain a reducible configuration in that n-range.

5. Possible improvements

While working out the discharging procedure and the unavoidable set $\mathscr{U}$ presented in this paper, we have been guided quite effectively by the preceding probability considerations.

We have always refused to accept any 15-ring configurations as members of the unavoidable set. Whenever we had no better choice than a 15- (or greater-) ring configuration we have changed the discharging procedure (by changing the sets of T-, S-, and L-discharging situations). In most cases, these changes also seemed to reduce the number of configurations in the unavoidable set and to simplify the argument. In some cases, however, we may have accepted more than ten configurations with $n < 15$ in order to replace one with $n = 15$.

On the other hand, whenever a configuration with $n \leq 14$ arose, if we found it "easily machine reducible" (i.e., D-reducible in a reasonable length of time or C-reducible with a reducer of the type accessible to our programs), we accepted it. We are aware of the fact that in many cases a more careful argument (without even changing the discharging procedure) might have shown that the configuration was not actually required (but replaceable by others already accepted). See footnote 4 on page 460.

In this way we have produced an argument which is definitely not best possible but, in our opinion, "reasonably close for a first try" to the best possible.

At this point we remark that one may try to "improve" the argument in quite different directions. For instance, one may try to absolutely minimize the number of configurations in the unavoidable set (possibly at the cost of a more complicated discharging procedure, of larger ring sizes, and of a greater amount of computer time needed). Or, one may try to minimize the ring-size of the configurations (possibly at the cost of increasing the number of configurations). Third, one may try to simplify as much as possible that part of the work which is to be done by hand (at the cost of increasing the work to be done by computer). Fourth, one may try to minimize the total combinatorial complexity of the argument, regardless of whether it is handled by computer or by hand.

It appears to be an interesting question how to reasonably define the total complexity of an argument of this kind. One might distinguish the logical complexity and the combinatorial complexity.

We believe that our argument is logically very simple and that all the com-

plexity is of the combinatorial nature. The relatively large number of about 485 individual S- and L-discharging situations makes the logically simple procedure tedious to carry out, since a large number of cases must be considered. The resulting large number of configurations in the unavoidable set $\mathcal{U}$ requires the same large number of reducibility proofs. Again, every reducibility proof is logically quite simple but requires the treatment of a great many different colorations, and for many of these colorations the treatment of a large number of Kempe chain arguments is required.

One might define the combinatorial complexity of a reduction proof (for an individual configuration C) to be the sum of the number of those good_0 colorations which are required and of all of the required colorations which are good_k ($k > 0$), each of these multiplied by the number of different Kempe interchanges which are required in order to relate it back to good_{k-1} colorations (in all possible cases of Kempe chain dispositions).

Applying this concept, the proof of D-reducibility of a 13-ring configuration will have a complexity of well over 10^6, while a proof of C-reducibility using a particularly suitable reducer may be of complexity under 10^4 and in extremely nice cases even under 10^3. Although present-day computers are perfectly capable of computing D-reduction, one might consider letting the computer search for the best possible C-reducer it can find (even if the configuration is D-reducible) and then choose the shortest proof of C-reducibility it can find. In many cases, such a proof could be checked by hand with reasonable effort.

The total combinatorial complexity of the entire argument would then be the sum of the complexities of all the reduction proofs required (plus the number of case distinctions required for the proof of unavoidability, which, however, can be expected to be much smaller than the complexity of the reduction proofs). We think that it would be interesting to obtain a reasonably good estimate of the minimum combinatorial complexity which is required for a proof of the Four Color Theorem.

The choice of a discharging procedure may be viewed as the consequence of a sequence of major decisions which must be made at certain stages of an otherwise routine process. We begin with the extremely simple procedure $\mathscr{P}(\emptyset, \emptyset, \emptyset)$ (see Section 3) and we call the corresponding charge function q_R. (q_R is obtained from q_0 by "regular" discharging of 30 along all 5-to-major edges.) This yields a relatively small list of configurations called q_R-positive. Some of these contain reducible sub-configurations which become members of the unavoidable set $\widetilde{\mathcal{U}}$ which will be constructed by the procedure. The remaining q_R-positive configurations are called critical. If we have accepted a reducible subconfiguration for $\widetilde{\mathcal{U}}$ from every q_R-positive configuration which contains one, the critical configurations at this point will be $CTS\#\#01$, 02, 03 (the configurations from Table 3 with arrows deleted), and

and $CTL\#\#1, \ldots, 22, 28, \ldots, 36, 81, \ldots, 90$, and 141 (from Table 4 with arrows deleted).

At this point the first major decision is made; we must choose the "long range dischargings" to use in $\mathscr{P}$. In this paper, we have chosen the set $\mathscr{T}$ of seven T-discharging situations (Figure 2); but one might try other choices. Now the discharging procedure $\mathscr{P}(\mathscr{T}, \emptyset, \emptyset)$ and the corresponding charge function q_T have been defined. Once $\mathscr{T}$ has been chosen in this manner, the two q_R-positive configurations drawn above are no longer q_T-positive; but on the other hand we must add to the collection of critical q_T-positive configurations the configurations $CTL\#\#11, 12$ (with arrows), $23, \ldots, 27, 37, \ldots, 55, 91, \ldots, 94$ (from Table 4).

Now our second major decision is the choice of small-discharging situations to use in order to avoid positive charges at the major vertices V_k in the critical $[q_T(V_k) > 0, k \geq 7]$-situations above. This decision is made by specifying a set $\mathscr{S}_0$ of "primary S-discharging situations". (In this paper about 70 members of Table 1 are primary in this sense.) The choice of $\mathscr{S}_0$ yields the discharging procedure $\mathscr{P}(\mathscr{T}, \mathscr{S}_0, \emptyset)$ and the corresponding charge function q_{TS_0}. Of course, we shall have made a basic decision as to which reducible configurations should be admitted to $\widetilde{\mathscr{U}}$. (In this paper, no 15-ring configurations were admitted. One might instead admit only configurations which are easily reducible in some sense; but this decision depends on the goal one has in mind.)

Next, it is a purely mechanical procedure to enumerate all critical q_{TS_0}-positive configurations. Then we must decide which large-discharging situations to use in order to avoid positive charges at the central V_5's of these critical configurations. In most cases, this decision will be "automatic". (In the procedure we used it was automatic in *all* cases.) This happens because at most one V_5-to-major edge which leaves a q_{TS_0}-positive V_5 will *not* be an S_0-edge, and such an edge is the natural choice for an L-discharging edge. The set $\mathscr{L}_0$ of L-discharging situations chosen yields $\mathscr{P}(\mathscr{T}, \mathscr{S}_0, \mathscr{L}_0)$ and $q_{TS_0L_0}$. At this point, additional S-situations must be chosen in order to take care of the critical $q_{TS_0L_0}$-positive major vertices. This yields an enlarged set $\mathscr{S}_1$ ($\mathscr{S}_0 \subset \mathscr{S}_1$), $\mathscr{P}(\mathscr{T}, \mathscr{S}_1, \mathscr{L}_0)$, and $q_{TS_1L_0}$. This process iterates until, at some stage, no critical positive situations remain. Then construction of the discharging procedure and the unavoidable set is finished. (For the discharging procedure presented in this paper we needed three stages of additional S- and L-situations $\mathscr{S} = \mathscr{S}_3$, $\mathscr{L} = \mathscr{L}_3$ where $\mathscr{L}_3$ was obtained from $\mathscr{L}_2$ by adding only two members.)

It is interesting to ask why this process "must" terminate, provided that the decisions are "reasonably" made. The answer can be obtained from the observation that from stage to stage the critical positive configurations contain more and more non-5 neighbors of the positive central vertex.

We did not carry out any further experiments of the type described above since the purely mechanical enumeration processes for the critical positive configurations are quite time-consuming if done by hand. Of course, since the enumeration processes are logically very simple, it is possible to write a com-

puter program to carry them out. Then, the work to be done by hand would consist of making the major decisions described above and would not involve much labor. But at such a stage of automation, how should one define an improvement of the procedure?

At the present stage of the development, many mathematicians might prefer a method by which the reduction proofs are done by machine but everything else is done by hand. Probably the best way to achieve this end would be to exhaustively compute 12- and 13-rings (in the same manner that Allaire and Swart treated 10-rings in [2] and 11-rings in their not yet published work). This would take care of the most tedious part of the combinatorial work. It is much easier to generate a configuration than it is to check whether this con-figuration contains a sub-configuration belonging to a given large set $\mathcal{U}$ of reducible configurations. It is easy, however, to verify that a configuration contains a sub-configuration with $n \leq 13$ and say $\phi \leq 1$. If this is done, only the few irreducible configurations of $n \leq 13$ and the required reducible con-figurations with $n = 14$ need be listed. (And the number of the latter could certainly be kept much smaller than the approximately 660 in our set $\mathcal{U}$.)

BIBLIOGRAPHY

1. F. ALLAIRE, *A minimal 5-chromatic planar graph contains a 6-valent vertex*, to appear.
2. F. ALLAIRE AND E. R. SWART, *A systematic approach to the determination of reducible configurations in the four-color conjecture*, J. Combinatorial Theory (B), to appear.
3. K. APPEL AND W. HAKEN, *An unavoidable set of configurations in planar triangulations*, J. Combinatorial Theory (B), to appear.
4. ———, *The existence of unavoidable sets of geographically good configurations*, Illinois J. Math., vol. 20 (1976), pp. 218–297.
5. K. APPEL, W. HAKEN, AND J. MAYER, *Triangulations à v_5 séparés dans le problème des quatre couleurs*, J. Combinatorial Theory (B), to appear.
6. A. BERNHART, *Six rings in minimal five color maps*, Amer. J. Math., vol. 69 (1947), pp. 391–412.
7. ———, *Another reducible edge configuration*, Amer. J. Math., vol. 70 (1948), pp. 144–146.
8. F. BERNHART, *On the characterization of reductions of small order*, J. Combinatorial Theory (B), to appear.
9. F. BERNHART AND S. GILL, *An extension of Winn's result on reducible minor neighborhoods*, to appear.
10. G. D. BIRKHOFF, *The reducibility of maps*, Amer. J. Math., vol. 35 (1913), pp. 114–128.
11. C. CHOJNACKI, *A contribution to the four color problem*, Amer. J. Math., vol. 64 (1942), pp. 36–54.
12. K. DÜRRE, *Untersuchungen an Mengen von Signierungen*, Doctoral Dissertation, Tech-nische Universität Hannover, 1969.
13. P. ERRERA, *Une contribution au problème des quatre couleurs*, Bull. Soc. Math. France, vol. 53 (1925), pp. 42–55.
14. PH. FRANKLIN, *The four color problem*, Amer. J. Math., vol. 44 (1922), pp. 225–236.
15. W. HAKEN, *An existence theorem for planar maps*, J. Combinatorial Theory, vol. 14 (1973), pp. 180–184.
16. H. HEESCH, *Untersuchungen zum Vierfarbenproblem*, B-I-Hochschulscripten 810/810a/810b, Bibliographisches Institut, Mannheim/Vienna/Zurich, 1969.
17. ———, *Chromatic reduction of the triangulations T_e, $e = e_5 + e_7$*, J. Combinatorial Theory, vol. 13 (1972), pp. 46–53.

18. P. J. HEAWOOD, *Map color theorem*, Quart. J. Pure Appl. Math., vol. 24 (1890), pp. 332–338.

19. A. B. KEMPE, *On the geographical problem of the four colors*, Amer. J. Math., vol. 2 (1879), pp. 193–200.

20. J. KOCH, *Computation of four color reducibility*, Ph.D. Thesis, 1976, Department of Computer Science Report no. UIUCDCS-R-76-802, University of Illinois at Urbana-Champaign, Urbana, Illinois.

21. H. LEBESGUE, *Quelques conséquences simples de la Formule d'Euler*, J. de Math., vol. 9 (1940), pp. 27–43.

22. J. MAYER, *Problème des quatre couleurs, un contre-exemple doit avoir au moins 96 sommets*, J. Combinatorial Theory (B), to appear.

23. O. ORE AND J. STEMPLE, *Numerical calculations on the four color problem*, J. Combinatorial Theory, vol. 8 (1970), pp. 65–78.

24. T. OSGOOD, *An existence theorem for planar triangulations with vertices of degree five, six, and eight*, Ph.D. Thesis, 1974, University of Illinois at Urbana-Champaign, Urbana, Illinois.

25. R. STANIK, *Zur Reduction von Triangulationen*, Doctoral Dissertation, Technische Universität Hannover, 1974.

26. W. STROMQUIST, *Some aspects of the four color problem*, Ph.D. Thesis, Harvard University, 1975.

27. W. TUTTE AND H. WHITNEY, *Kempe chains and the four colour problem*, Utilitas Mathematica, vol. 2 (1972), pp. 241–281, reprinted in Studies in Graph Theory, Studies in Math., Math. Assoc. of America, vol. 12, 1976.

28. P. WERNICKE, *Über den kartographischen Vierfarbensatz*, Math. Ann., vol. 58 (1904), pp. 413–426.

29. C. WINN, *On the minimum number of polygons in an irreducible map*, Amer. J. Math., vol. 62 (1940), pp. 406–416.

Added in proof. On September 21, 1977 we received the following preprint containing the list of 2669 reductions referred to on page 491 of this journal.

30. K. DÜRRE, H. HEESCH, AND F. MIEHE, *Eine Figurenliste zur Chromatischen Reduktion*, Preprint No. 73 (1977), Institut für Mathematik der TU Hannover.

On this list we found 415 configurations which are members of our set $\mathscr{U}'$ of which 256 are D-reducible. We also found reducible sub-configurations of 227 of the members of $\mathscr{U}'$. (These latter permit a decrease in the size of $\mathscr{U}'$ by 77, thus the smallest current unavoidable set has 1405 members.)

In all cases of configurations appearing on both lists there is agreement on whether or not the configuration is D-reducible. For those that are not D-reducible there is exact agreement on the number of bad colorings after the D-algorithm is applied (these numbers are given in the microfiche supplement) with the exception of two articulated 11-ring configurations for which the list in [30] shows fewer bad colorings. In these two cases our results agree with those of Allaire's program.

UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS