

On the Dissection of Rectangles into Right-Angled Isosceles Triangles

J. D. Skinner II

1740 North 59th Street, Lincoln, Nebraska 68505-1123

C. A. B. Smith

141 Portland Crescent, Stanmore HA7 1LR, England

and

W. T. Tutte

*Department of Combinatorics and Optimization, University of Waterloo,
Waterloo, Ontario, N2L 3G1 Canada*

Received July 2, 1999

We consider the problem of dissecting a rectangle or a square into unequal right-angled isosceles triangles. This is regarded as a generalization of the well-known and much-solved problem of dissecting such figures into unequal squares. There is an analogous “electrical” theory but it is based on digraphs instead of graphs and has an appropriate modification of Kirchhoff’s first law. The operation of reversing all edges in the digraph is found to be of great help in the construction of “perfect” dissected squares. © 2000 Academic Press

1. INTRODUCTION

This paper is an investigation into two related matters:

- (i) The dissection of rectangles into isosceles right-angled triangles.
- (ii) Generalized electrical networks

The *generalized* electricity is of the kind called *leaky* in [3] and *unsymmetrical* in [7].

We require the concepts of *edge* and *vertex* as related to polygons in the plane and as graph-theoretical concepts. To avoid confusion we will write, usually, of *sides* and *corners* of polygons and of *darts* (or directed edges)

and *nodes* of digraphs. A polygon is deemed to include its interior and a side its end-corners.

By a *dissection* Δ of a polygon Π we mean a finite set of polygons Π_r , called *elements* or *tiles*, whose union is Π and which meet, if at all, only in sides and corners. An *RI-dissection* is a dissection of a rectangle R whose tiles are right-angled isosceles triangles. This investigation of RI-dissections follows naturally on two others, of the dissection of rectangles into squares [2] and of the dissection of equilateral triangles, or 60° parallelograms, into equilateral triangles [6].

The authors of [2] show a fleeting interest in RI-dissections in the last sentence of their paragraph 10.3. Interest was revived recently by one of us (J.D.S.) with his *perfect* RI-dissections of squares.

2. RI-DISSECTIONS AND S-DISSECTIONS

Without loss of generality we suppose the sides of R to be horizontal and vertical. If the tiles are all of different sizes we say that the RI-dissection is *perfect*. Figure 1A shows a perfect RI-dissection.

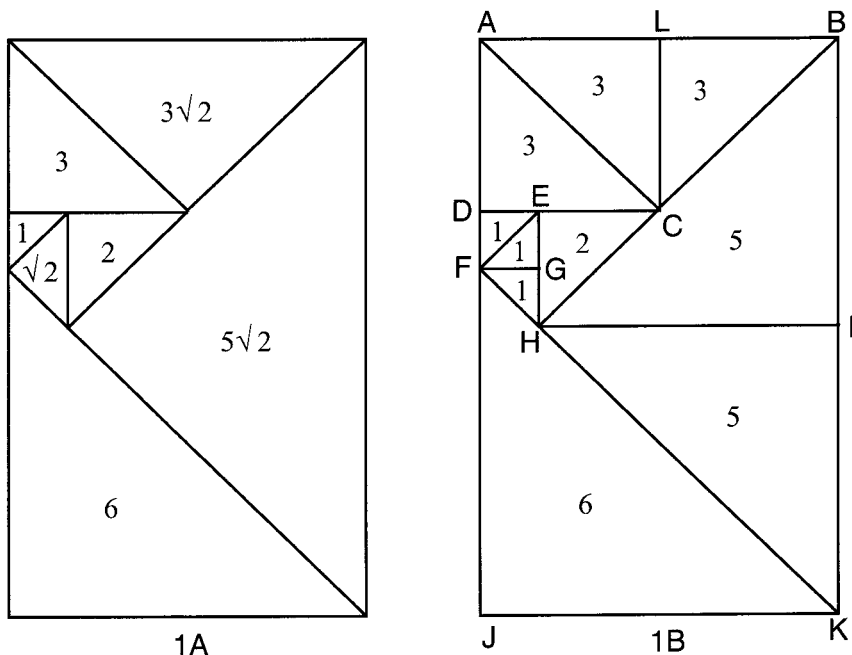


FIG. 1. From RI to S.

It is easy to verify that each tile has sides inclined to the horizontal at multiples of 45° , so that a tile has eight possible orientations. Each tile has either (a) one horizontal, one vertical, and one sloping side or (b) two sloping sides and the other either horizontal or vertical. If every tile is of type (a) we call the dissection Δ a *special dissection* or an *S-dissection*. In the theory of this paper, our concern is mainly with S-dissections, though we point out in Section 10 that they can be used as a first stage in the construction of perfect RI-dissections. Any RI-dissection can be changed into an S-dissection by bisecting each tile of type (b) by a line through its right-angled corner. Thus the RI-dissection of Fig. 1A gives rise to the S-dissection of Fig. 1B. It can be shown that no S-dissection is perfect.

A dissection of a rectangle into squares becomes an S-dissection when each tile is bisected along a diagonal. A dissection into equilateral triangles becomes an S-dissection under a horizontal shear, followed by a slight vertical expansion. So the two earlier investigations can be regarded as special cases of this one.

For a tile T of an S-dissection it is often convenient to call its horizontal side its *base*, its vertical side its *cobase*, and its sloping side its *hypotenuse*. The corner opposite the base is the *apex* of T and the one opposite the cobase is the *coapex*. T is *up-pointing* or *down-pointing* when the apex is above or below the base respectively. It is *left-pointing* or *right-pointing* when as the coapex is to the left or right of the cobase respectively.

Figure 2 shows part of an S-dissection. In it tile ABE had base BE , cobase AE , apex A , and coapex B . It is up-pointing and left-pointing. Tile CGH is down-pointing and left-pointing and CDE is up-pointing and right-pointing.

In the remainder of this section we study the properties of an S-dissection Δ of a rectangle R . We define the *frame* of Δ as the union of the sides

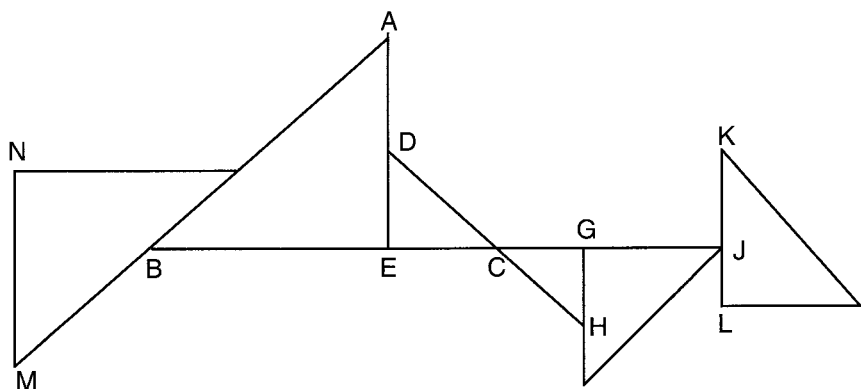


FIG. 2. Part of an S-dissection.

of its tiles. A *frame-segment* of Δ is a straight segment contained in the frame. It is said to *traverse* its internal points but not its two ends. A *maximal segment* of Δ is a frame-segment that cannot be extended in either direction within the frame. Thus tile-sides are frame-segments and the sides of R are maximal segments. Evidently each frame-segment extends uniquely as a maximal segment. In Fig. 2 BG is a frame-segment and BJ and DH are maximal segments.

A point P traversed by two distinct maximal segments in a *cross* of Δ , for example C in Fig. 3. A cross is necessarily a tile-corner.

THEOREM 2.1. *Let P be a tile-corner. Let L be a half-line with end P , drawn in a horizontal or vertical direction. Then one of the following propositions holds.*

- (i) L contains a tile-side with end P .
- (ii) L lies, except for P , entirely outside R .
- (iii) L passes through a tile with a side traversing P .

Proof. By symmetry we need only discuss the case in which L is drawn upward from P . Then if the theorem fails, L passes through a tile T incident with P and having an angle of 45° or 90° there. At least one arm of this angle must be vertical or horizontal. But this is absurd; each arm must make a non-zero angle of less than 90° with L . ■

THEOREM 2.2. *A tile corner P of Δ is a cross if and only if it is inside R and is not traversed by any tile-side. Moreover, each cross is traversed by a maximal horizontal and a maximal vertical segment.*

Proof. Clearly the conditions specified hold at every cross. Whenever P satisfies them, Theorem 2.1 tells us that all four of the possible half-lines L satisfy proposition (i). The theorem follows. ■

When Δ has crosses it is convenient to subdivide some of the maximal segments into smaller segments, called *effective segments* or *e-segments*, by cutting them at some, but not necessarily all, of the crosses they traverse. Maximal segments not cut at all in this process are also counted as e-segments. We try to make the cuts so that each cross is traversed by just one e-segment. We arrange this by cutting either the vertical or the horizontal segment, but not both, at each cross Q and then cutting at Q each sloping segment that traverses it. So no sloping e-segment traverses a cross. From now on we suppose such a set of e-segments defined. Since the sides of R traverse no crosses they are e-segments. Each tile-side is in an e-segment, by 2.2.

THEOREM 2.3. *Except for the four corners of R each tile-corner P is traversed by exactly one e-segment.*

Proof. If P is on a side of R the result is obvious. If P is a cross, it is a consequence of the definition of an e-segment. In the remaining case it is traversed by one e-segment, by 2.2, but not by more since e-segments do not cross. ■

3. PROPERTIES OF e-SEGMENTS

We consider a S-dissection Δ of a rectangle R .

We note first that there are two kinds of sloping e-segments, up-sloping and down-sloping when the segment rises or falls as we follow it from left to right in R , respectively.

Let AB be an e-segment of Δ that is not a side of R . Its ends A and B are traversed by uniquely determined e-segments $E(A)$ and $E(B)$, respectively, by 2.3. Each is differently oriented from AB since two e-segments have at most one point in common. We call them the *enclosers* of AB at A and B , respectively.

In Fig. 1B there are no crosses and so the e-segments are simply the maximal segments. There HI has enclosers FK and BK . The enclosers of EH are CD and FK . In Fig. 2, with CDE a tile, CD must be an e-segment since C is a cross. One of its enclosers contains AE and the other is the vertical or horizontal e-segment that traverses C .

We extend the definition by defining the enclosers of a side of R as the two adjacent sides.

The two sides of an e-segment AB , that is the two half-planes separated by its line, are naturally distinguished as upper and lower if AB is horizontal and as left and right if AB is vertical. If AB is sloping, either distinction is acceptable.

If AB traverses a tile-corner C we say that C is *active* on the side Z of AB if it is the end of an e-segment lying (except for C) in Z . Thus C is active on at least one side of AB , any side containing a tile of which C is a corner. If C is active on both sides it is a cross, by 2.2.

THEOREM 3.1. *Let an e-segment AB of Δ traverse a tile-corner C . Let C be active on the side Z of AB . Then, counting AB itself, there are at C on the side Z just one horizontal e-segment, just one vertical one, and at most two sloping ones. If there are two sloping ones then one is up-sloping and the other down-sloping.*

Proof. If AB is horizontal there is a vertical e-segment CD at C in Z , by 2.1. Each of the right angles ACD and BCD may or may not be bisected

by a sloping e-segment. If both are so bisected the two sloping e-segments are at right angles. We argue similarly if AB is vertical.

If AB is sloping there is a horizontal and a vertical e-segment at C in Z , by 2.1. There is room for just one more e-segment at C in Z , a sloping one at right angles to AB . ■

In view of Theorem 2.3 this result gives us information about all the tile-corners of Δ , except for the four corners of R . However, it is clear that at a corner of R there is just one horizontal e-segment, just one vertical, and at most one sloping.

The theorem can be exemplified by the tile-corners of Fig. 1B, say G with no sloping e-segments, E with just one, and C with two.

Let AB be an e-segment and T a tile. By the definition of e-segments the intersection of T with AB is a side of T , a corner of T , or null. In the first case we say that T is *side-incident* with AB .

We say T is *corner-incident* with AB if it has only a corner X in AB and T is not separated from AB by an encloser of AB . Such a separation can occur only if X is A or B and then only if X is a cross, by 2.2.

Let Z be a side of AB on which there is an incident tile. Consider the sequence

$$P = (A_0, A_1, A_2, \dots, A_K) \quad (1)$$

of tile corners on AB that is defined as follows. P starts with $A_0 = A$, continues with the tile corners traversed by AB that are active on the side Z , in their natural order from A to B , and ends with $A_K = B$. We call P the *Z-side corner-sequence* of AB from A to B .

If A_i and A_{i+1} are consecutive members of P , then the segment $A_i A_{i+1}$ is a side of tile T_0 that lies on the Z -side of AB and is side-incident with AB . We call the sequence

$$Q = (T_0, T_1, \dots, T_{K-1}) \quad (2)$$

the *Z-side tile-sequence* of AB from A to B .

Since tiles do not overlap we can assert the following theorem.

THEOREM 3.2. *A tile T is side-incident with an e-segment AB and lies on its side Z if and only if it belongs to the Z-side tile-sequence of AB from A to B .*

In particular cases the term *Z-side* can be replaced in an obvious way by *upper*, *lower*, *left*, or *right*.

From obvious geometrical considerations we can assert

THEOREM 3.3. *Let Ab be a sloping e-segment and T a tile side-incident with it. If AB is up-sloping then if T is on its upper side T is down-pointing*

and right-pointing, but if T is on its lower side T is up-pointing and left-pointing. Similarly if AB is down-sloping then T is down-pointing and left-pointing if on the upper side, but up-pointing and right-pointing if on the lower side.

THEOREM 3.4. *Let AB be a horizontal or vertical e-segment. Then the number of tiles side-incident with it is equal to the number corner-incident with it.*

Proof. It is enough to consider the case in which AB is horizontal. (See below.) Suppose first that AB is the lower horizontal side of R . Let its upper corner-sequence from A to B be given by (1) and the corresponding tile-sequence by (2). At each member A_i of (1) there is a vertical e-segment E_i , by 2.1. If A_i and A_{i+1} are consecutive members of (1) we have between them the tile T_i with base A_iA_{i+1} , which is side-incident with AB . There is also a tile filling the angle between T_i and one of the two e-segments, E_i if T_i is left-pointing and E_{i+1} if T_i is right-pointing. This tile is corner-incident with AB . There is no room for a third tile incident in any way with AB between E_i and E_{i+1} . The theorem follows for this case.

For the upper horizontal side of R the argument is similar, with the E_i now below AB , not above.

The remaining case is more complicated since the enclosers need not be vertical. Take A to be the left end and B the right. At each tile corner traversed by AB and active above (below) AB there is a vertical e-segment extending above (below) AB . If the encloser at A or B is vertical it serves as the vertical e-segment there on both the upper and the lower sides. If it is down-sloping at A or up-sloping at B the vertical e-segment there extends upwards. In the remaining case, it extends downwards. Taking these vertical e-segments in clockwise order around AB we still find just one side-incident tile and just one corner-incident tile between any two consecutive ones. This completes the proof. ■

By saying that for 3.4 we need only consider the horizontal case we mean that we can then deduce the vertical case by turning R through 90° and applying the horizontal case. General theorems about S-dissections come in dual pairs, related by this device.

4. SIZES AND CO-ORDINATES

We define the *size* of a tile in the S-dissection Δ of R as the length of its horizontal or vertical side with respect to some agreed unit of measurement. We write H and V for the lengths of the horizontal and vertical sides of R , respectively. For the length of a segment AB we write $L(AB)$.

Let AB be a horizontal or vertical e -segment. Considering the tile sequence Q from A to B on one of its sides, we see that the sum of the sizes of the members of Q is $L(AB)$. So, by 3.2 we have

THEOREM 4.1. *Let AB be a horizontal or vertical e -segment and Z a side of AB not exterior to R . Then $L(AB)$ is the sum of the sizes of the tiles side-incident with AB on the side Z .*

We take the lower left corner of R as the origin of a system of rectangular Cartesian coordinates, the x -axis proceeding along the lower horizontal side and the y -axis along the left vertical side. The y -coordinate of a point or horizontal segment K is the *potential* $P(K)$ of K . The x -coordinate of a point or vertical segment K is the *copotential* $P^*(K)$ of K . Thus the upper and lower horizontal sides of R have potentials V and 0 , respectively, while the left and right vertical sides have copotentials 0 and H , respectively.

Imagine that the rectangle is made of electrically conductive material of unit conductivity and that a uniform current is passed through it from top to bottom. Then this current will enter a down-pointing tile at its base (top) and "leak away" until the current becomes zero at the apex (bottom). The opposite will happen in an up-pointing tile. These considerations lead to the following definitions.

The *current* $i(T)$ in a tile T is its size if T is down-pointing and minus the size if T is up-pointing. The *cocurrent* $j(T)$ in T is its size if T is left-pointing and minus the size if T is right-pointing. From the above definitions we have the following.

THEOREM 4.2. *Let ABC be a tile with base AB and cobase AC . Then*

$$i(ABC) = P(AB) - P(C) \quad \text{and}$$

$$j(ABC) = P^*(AC) - P^*(B).$$

For the same tile one can by inspection verify the following

THEOREM 4.3. *If BC is up-sloping then $i(ABC)$ and $j(ABC)$ have opposite signs. If BC is down-sloping they have the same sign.*

From 4.1 and 4.2 we can deduce

THEOREM 4.4. *The sum of the currents $i(T)$ in the side-incident tiles is H for the upper horizontal side of R and $-H$ for the lower one. It is zero for any other horizontal e -segment.*

Dually the sum of the $j(T)$ over the side-incident tiles is V for the right vertical side of R and $-V$ for the left one. It is zero for any other vertical e-segment.

5. REPRESENTATIONS BY GRAPHS AND MAPS

In [2] and [6] the authors hasten to replace their dissections by graphs or digraphs, structures that look simpler and are easier to draw, but contain the same mathematical information. We now attempt a similar replacement.

Imagine an e-segment AB of the S-dissection Δ of R to be slit from A to B . Into the slit we insert a topological disk, identifying one half-circumference with one side of the slit and the complementary half-circumference with the other. Less formally we can speak of pulling the sides of the slit a little apart and filling the gap with a lune. Topologically we still have a diagram on a plane, even on a sphere if we take, as we shall, the plane to be closed by a point Ω at infinity.

We apply this operation to every e-segment. Since e-segments do not cross there is no mutual interference. We thus get a spherical map $M_1(\Delta)$, the "first derived map of Δ ." Its faces are the slit-fillers, the tiles of Δ and the exterior of R . Its vertices are the non-cross tile-corners of Δ together with two vertices for each cross. For the slitting pulls each cross apart into two vertices of $M_1(\Delta)$, one on each side of the slit.

We now partition each tile T into three triangles by lines drawn from an interior point, called the *centre* of T , to the three corners. We annex each of these triangles to the slit-filler with which it has a common side, to make a single region. We partition the exterior of R into four regions by segments from the four corners of R to Ω , each at 135° to the adjacent sides of R . We annex each new region to the slit-filler of the incident side of R . We now have a new spherical map $M_2(\Delta)$, the *second derived map* of Δ .

The faces of $M_2(\Delta)$ are expanded slit-fillers, classed as horizontal, vertical, up-sloping, or down-sloping according to the nature of the enclosed e-segment. The four faces incident with Ω are the upper, lower, left, and right *outsiders*, according to which side of R is enclosed.

The nodes of $M_2(\Delta)$ are of three kinds, the *outer* node Ω , the *central* nodes or tile-centres, and the nodes of $M_1(\Delta)$ which we now call *corner-nodes*.

Each edge of $M_2(\Delta)$ has one end Ω or a central node and as the other a corner-node.

As an example take the Δ of Fig. 1B. To visualize $M_1(\Delta)$ imagine each e-segment of that figure slit and gapping for its lune. The second derived

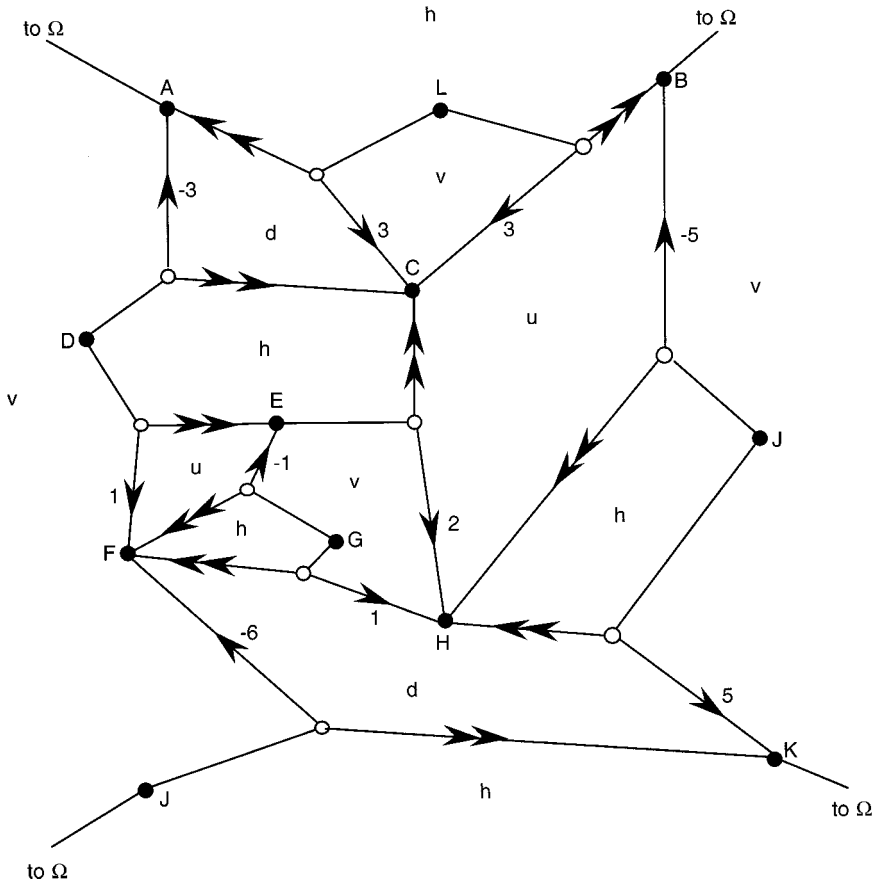


FIG. 3. Second derived map of Fig. 1B.

map is shown in Fig. 3 where the corner-vertices are labelled as in Fig. 1B. The letters h , v , u , and d indicate horizontal, vertical, up-sloping, and down-sloping faces, respectively. The arrows and numbers indicate darts and currents as defined below.

THEOREM 5.1. *The second derived map of an S -dissection Δ has the following properties.*

- (i) *It is bipartite, one side of the bipartition having the corner-nodes, the other Ω and the central nodes.*
- (ii) *Ω is incident only with the four outsiders, these being alternately horizontal and vertical here.*

(iii) *Each central node is trivalent, incident with one horizontal, one vertical, and one sloping face.*

(iv) *Each corner-node is at most tetravalent, incident with just one horizontal, just one vertical, and at most two sloping faces.*

(v) *The boundary of each face is a circuit. The intersection of two face-boundaries, if non-null, is connected, being a single node, an arc of length 1, or an arc of length 2 with a corner-node as internal vertex.*

Proof. (i), (ii), and (iii) are immediate consequences of the construction and (iv) follows from 3.1 Proposition (v) follows from the fact that two e-segments meet in at most one point and that central vertices are trivalent. ■

Let E be an edge of $M_2(\Delta)$ incident with a central node C , but which is not a side of the horizontal face F at C . The other end of E is a corner-vertex B at which there is just one horizontal face G . We note that by the construction of $M_2(\Delta)$, C represents a tile-side incident with the e-segment in F and corner-incident with that in G . We direct E from C to B , calling it from now on a dart. We say it is *from F to G , outgoing from F , and incoming to G .*

Dually we can define darts from one vertical face to another. But these we call *codarts*. Darts and codarts are marked with single and double arrows, respectively, in the diagrams.

THEOREM 5.2. *The system of darts and codarts of $M_2(\Delta)$ has the following properties.*

(i) *The number of darts outgoing from a given horizontal face is equal to the number incoming to it. And dually for a vertical face and its associated codarts.*

(ii) *The darts and codarts make up the bounding circuits of the sloping faces.*

(iii) *Darts and codarts alternate in the boundary of any sloping face, darts going one way round the face and codarts the other.*

Proof. Proposition (i) follows from 3.4.

The dart and codart directed from a central vertex C are the only edges at C incident with a sloping face. (5.1, (iii)). This implies (ii).

The edges of a sloping face F are incident alternately with horizontal and vertical faces (5.1, (iii) and (iv)). Hence the edges are alternately darts and codarts. At each central node incident with F the dart and codart are oppositely directed around F . Proposition (iii) follows. ■

We must now take account of potentials and currents. To each horizontal or vertical face we assign the potential or copotential, respectively, of the corresponding e-segment. To the dart D directed from a central node C we assign the current $i(T)$ in the tile T with centre C . It is the fall of potential from the horizontal face F incident with C to that incident with the other end of D . Dually we assign the cocurrent $j(T)$ to the codart from C .

The dart D and the codart E from C are incident with a sloping face Q . Suppose Q up-sloping. Then, by 3.3, T is down-pointing and right-pointing if above the e-segment in Q and up-pointing and left-pointing if below it. In either case the dart goes counter-clockwise around Q and the codart clockwise. By this and the analogous argument with Q down-sloping we have

THEOREM 5.3. *Darts go counter-clockwise and codarts clockwise around an up-sloping face, but clockwise and counter-clockwise, respectively, around a down-sloping one.*

COROLLARY 5.3.1. *If a corner-vertex K has two incident sloping faces then one is up-sloping and the other down-sloping.*

This is because no dart is directed to K .

By 4.3 and 4.4 we have the following rules of currents and cocurrents.

THEOREM 5.4. (i) *The sum of the currents in the darts from a horizontal face F is H for the upper outsider, $-H$ for the lower one, and zero for each other F . There is a dual rule for vertical faces and cocurrents.*

(ii) *At each central node C the current $i(C)$ and the cocurrent $j(C)$ satisfy $i(C) = j(C)$ if the incident sloping face is down-sloping and $i(C) + j(C) = 0$ if it is upsloping.*

The electrical net $N(\Delta)$ of Δ can be defined as part of $M_2(\Delta)$. It is the diagraph whose nodes are the horizontal faces of $M_2(\Delta)$ and whose directed edges are the darts of $M_2(\Delta)$. Similarly the vertical faces and codarts define a dual electrical net $N^*(\Delta)$.

Naturally in separate drawings of $N(\Delta)$ the nodes are contracted into points, the upper and lower outsiders being first pulled apart at Ω and the other two combined in a single outer face. $N(\Delta)$ is shown thus in Fig. 4 for the Δ of Fig. 1B.

THEOREM 5.5. *$N(\Delta)$ is connected, even if the vertical outsiders are regarded as detached at Ω .*

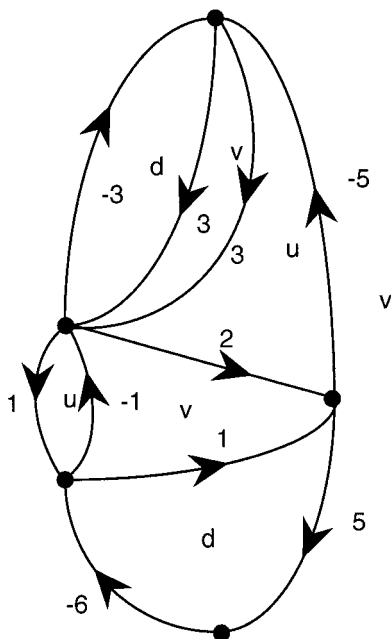


FIGURE 4

Proof. Suppose $N(\Delta)$ to have a component C not including the upper outsider (as a vertex). Let v_r be a node of C of highest potential. Then v_r corresponds to a horizontal e-segment with no side-incident tile on its upper side. Then that e-segment is the upper side of R , contrary to the choice of C . ■

6. UNSYMMETRICAL ELECTRICITY

This section derives from a study by one of us (C.A.B.S.) of the electrical theory in [2] and [6]. There is a more recent and more general treatment in [4].

A *dinet* X consists of a loopless digraph Γ together with a potential P_r associated with each node v_r and a conductance $c(D)$ associated with each dart D . If D is directed from v_r to v_s then the current in D is defined as the product of $c(D)$ and the potential difference $P_r - P_s$.

In addition two of the nodes are selected as the *source* and *sink*. The currents are required to satisfy Kirchhoff's (modified) first law, namely that at any node other than the source and sink the sum of the currents in the

darts directed away from it is zero. The dinet also trivially obeys Kirchhoff's second law: the sum of the potential differences around any circuit is zero.

As an example we can take the electrical net N of Section 5, assigning to each dart conductance 1. We use more general conductances here for the sake of a more general analogy with the usual theory of electrical networks. We can, if we wish, think of the conductances as complex numbers. Later we shall restrict them to being positive real numbers and later still to having the common value 1.

Let us suppose the suffix r of v_r and P_r to run from 1 to n . Let the source be v_a and the sink v_b . For each v_r we define c_{rr} as the sum of the conductances in the darts outgoing from v_r . For distinct nodes v_r and v_s we define c_{rs} as minus the sum of the conductances of the darts directed from v_r to v_s . Thus

$$\sum_s c_{rs} = 0. \quad (1)$$

We write the equations

$$\sum_s (-c_{rs})(P_r - P_s) = \sum_s c_{rs}P_s = I_r, \quad (r = 1, 2, \dots, n) \quad (2)$$

calling I_r the current *entering* X at v_r . By (1) we can assert the first law by saying that

$$I_r = 0, \quad (r \neq a, r \neq b). \quad (3)$$

We call the matrix $\{c_{rs}\}$ of Eqs. (2) the *Kirchhoff matrix* $K(X)$ of the dinet X . Its determinant is zero, by (1). Striking out its r th row and column we obtain the submatrix $K_r(X)$.

We may hope to solve (2) for the potentials, with I_a as a given quantity and I_b as another unknown to be determined. We can set $P_b = 0$. The equations other than the b th can be solved uniquely for the other potentials provided that $K_b(X)$ is non-singular. Then the b th equation determines I_b .

There is a solution of (2) that is given directly in terms of the structure of Γ . We define the *weight* $\omega(\Theta)$ of a subdigraph Θ of Γ as the product of the conductances of its darts. A subdigraph Θ may be a tree, in particular a tree directed to one of its nodes v_r , this meaning that each dart of Θ is directed towards v_r . All roads are one-way and lead to Rome, that is, to v_r .

The sum of the weights of all spanning trees of Γ directed to v_r is the *tree-sum* at v_r , denoted by $T_r(X)$. We are concerned also with spanning

double trees, spanning subdigraphs of Γ of two non-overlapping components, each of which is a tree directed to one of its vertices. For these it seems convenient to have symbols of the form

$$(acd \cdots j, bpq \cdots w).$$

This we define as the sum of the weights of all double trees in which one component is directed to v_a and contains also v_c, v_d, \dots, v_j , while the other is directed to v_b and contains also v_q, \dots, v_w . We also define the *transpedance* $[ab, xy]$ by

$$[ab, xy] = (ax, by) - (ay, bx). \tag{4}$$

THEOREM 6.1.

$$[ab, xy] + [ab, yz] = [ab, xz] \tag{5}$$

Proof.

$$\begin{aligned} [ab, xy] + [ab, yz] &= (ax, by) - (ay, bx) + (ay, bz) - (az, by) \\ &= \{(axz, by) + (ax, byz)\} - \{(ayz, bx) + (ay, bzx)\} \\ &\quad + \{(ayx, bz) + (ay, bxz)\} - \{(azx, by) + (az, bxy)\} \\ &= (ax, byz) - (ayz, bx) + (ayx, bz) - (az, bxy) \\ &= (ax, bz) - (az, bx) = [ab, xz] \quad \blacksquare \end{aligned}$$

We see from this that the transpedances define a set of potentials, $[ab, xy]$ being the fall of a potential from v_x to v_y when v_a is the source and v_b the sink. The fall $[ab, ab]$ from source to sink can be written also as (a, b) since (ab, ba) is an empty sum.

THEOREM 6.2.

$$\begin{aligned} \sum_D c(D)[ab, ax(D)] &= T_b(X) \quad \text{and} \\ \sum_D c(D)[ab, bx(D)] &= -T_a(X), \end{aligned}$$

where D runs through the set of darts outgoing from v_a in the first equation and from v_b in the second. In each case $x(D)$ is the suffix of the other end of D .

Proof. Considering the first equation we note that

$$[ab, ax(D)] = (aa, bx(D)) = (a, bx(D)),$$

since $(ax(D), ba)$ is an empty sum. So the term $c(D)[ab, ax(D)]$ is the weight of a spanning tree of Γ directed to v_b and having D as a dart. On the other hand the weight of any spanning tree of Γ directed to v_b can be written as $c(D)w(\Theta)$ where D is the dart of T outgoing from v_a and Θ is a double tree with components directed the one to v_a and the other to v_b . The first equation follows.

The proof of the second equation is similar, $[ab, bx(D)]$ being there equal to $-(ax(D), b)$. ■

THEOREM 6.3. *Let v_c be distinct from v_a and v_b . Then*

$$\sum_D c(D)[ab, cx(D)] = 0,$$

where D runs through the set of darts outgoing from v_c , and $x(D)$ is the suffix of the other end of D .

Proof. We rewrite the sum on the left as

$$\sum_D c(D)(ac, bx(D)) - \sum_D c(D)(ax(D), bc).$$

The typical term of the first sum derives from a double tree Q with two components A and B directed to v_a and v_b , respectively, with v_c in A and $v_{x(D)}$ in B . Now there is just one dart of A directed from v_c . Call it E . By deleting E from Q and adjoining D we get a double tree Q' with components A' and B' directed to v_a and v_b , respectively, and with v_c in B' and $v_{x(E)}$ in A' . (See Fig. 5.) The analogous operation on a typical term of the second sum deriving from a double tree Q' shows Q' to be derived from some Q in the above manner. But Q' contributes the same term to our second sum as does Q to the first. So the two sums are equal and the theorem is proved. ■

We see from this that the currents derived from the transpendances as potential differences satisfy Kirchhoff's first law. We shall refer to any set of currents and potentials satisfying this law as a *flow* from source to sink, and this particular flow we shall call the *full flow* from v_a to v_b .

The full flow is a solution of Eqs. (2) with $I_a = T_b(X)$, $I_b = -T_a(X)$ and $P_r = [ab, rb]$. If the matrix $K_b(X)$ is non-singular it is the only solution of those equations with $I_a = T_b(X)$.

There is a companion result to Theorem 6.1 obtained in the same way but with the roles of the first and second suffix-sets in $[pq, rs]$ interchanged. It is

$$[ab, xy] + [bc, xy] = [ac, xy]. \quad (6)$$

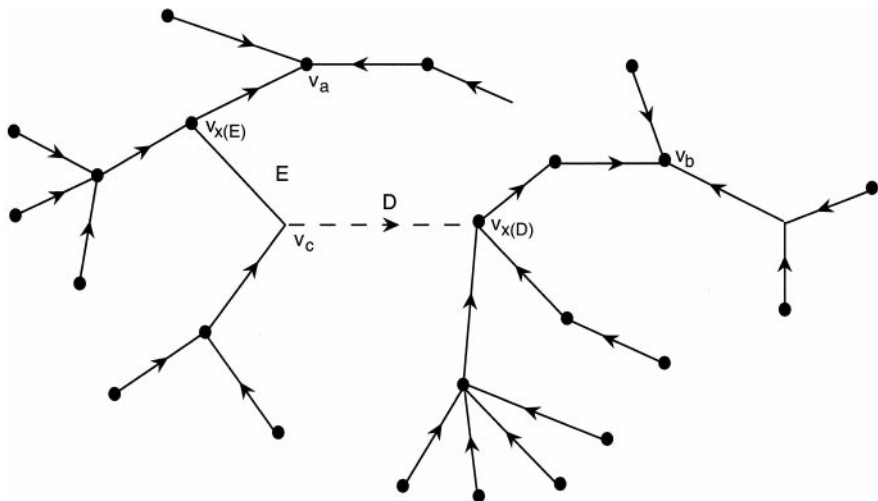


FIG. 5. A double tree.

In the standard theory of electrical networks, used in [2], this is not a new equation, for there we have the rule $[pq, rs] = [rs, pq]$. Perhaps we should emphasize that in the present theory that rule does not in general hold. Nor, in any flow of X , need the current I_x entering at the source be equal to the current $-I_b$ leaving at the sink.

There is a recursion formula for tree-sums. Suppose that in X there is a dart D directed from v_r to v_s . Deleting D we obtain a dinet X_D . Deleting all the darts joining v_r to v_s and then identifying those two nodes as a new node still denoted by v_s we get a dinet X^D . Of the spanning trees of X directed to v_s those not containing D are those of X_D . The others are in 1-1 correspondence with those of X^D . So we have

$$T_s(X) = T_s(X_D) + c(D) T_s(X^D). \quad (7)$$

There is a similar formula for the determinant of $K_s(X)$. The matrix $K_s(X)$ differs from $K_s(X_D)$ only by having an extra term $c(D)$ in the intersection of the row and column of v_r . And $K_s(X^D)$ is got from $K_s(X)$ by striking out the row and column of v_r . So by the theory of determinants

$$\det K_s(X) = \det K_s(X_D) + c(D) \det K_s(X^D). \quad (8)$$

THEOREM 6.4. For any node v_s is a dinet X

$$T_s(X) = \det K_s(X).$$

Proof. If Γ consists solely of v_s then $T_s(X) = 1$. $K_s(X)$ is then a matrix of order zero and its determinant is, by convention, 1. So we may suppose Γ to have at least two nodes.

If no dart of Γ is directed to v_s then $T_s(X) = 0$. Then in each row of $K_s(X)$ the elements sum to zero and so $\det K_s(X) = 0$.

We can complete the proof by induction over the number of vertices and the number of darts directed to v_s using the above preliminary results and the recursions (7) and (8), which are valid even when $K_s(X)$ has but one vertex. ■

This result is often called the matrix-tree theorem. We can apply it to Eqs. (2) as follows. Provided that $K_b(X)$ is non-singular we can solve the equations other than the b th for the full potentials by Cramer's rule, with $P_b = 0$. The quotient $T_b(X)/\det K_b(X)$ takes the value 1. Then P_r , that is $[ab, rb]$, is presented as the determinant of the submatrix Z of $K(X)$ obtained by striking out the a th and b th rows and the r th and b th columns, multiplied by the appropriate unit $+1$ or -1 . If $a > b$ and $r > b$ we can say that P_r is the cofactor, as defined in [1], of the complementary submatrix of Z . This means that the multiplier is 1 or -1 according as $a + r$ is even or odd. If just one of $a > b$ and $r > b$ is false an extra change of sign is required since

$$[pq, rs] = -[pq, sr] = -[qp, rs], \quad (9)$$

by the definition of a transpedance.

By the device of making not P_b but some other potential P_s zero we can similarly relate $P_r - P_s$, that is $[ab, rs]$ to the cofactor of $K(X)$ corresponding to the striking out of the a th and b th rows and the r th and s th columns.

7. BALANCED DINETS

The dinet X of Section 6 is said to be *balanced* if at each node the sum of the conductances of the outgoing darts is equal to that of the incoming ones. The electrical nets of Section 5 are balanced, by 5.1.

For a balanced dinet X we can complement 6.1 with

$$\sum_r c_{rs} = 0. \quad (1)$$

THEOREM 7.1. *In a balanced dinet X the tree sums $T_r(X)$ are all equal.*

Proof. For any choice of source v_a and sink v_b we sum the Eqs. 6.2, stated for the full flow from v_a to v_b . By (1) the result is

$$I_a + I_b = 0.$$

Hence, $T_a(X) = T_b(X)$. ■

We refer to the common value of the sums $T_r(X)$ as the tree-sum $T(X)$ of X . We now have a case in which the current I_a entering at v_a is equal to the current $-I_b$ leaving at v_b .

THEOREM 7.2. *Let S and T be complementary node-sets in a balanced dinet X . Then the sum of the conductances in the darts directed from S to T is equal to that in those directed from T to S .*

Proof. Let the first sum be \sum_1 and the second \sum_2 and let \sum_0 be the sum of the conductances of the darts from S to S . The sums over all nodes of S for outgoing and incoming darts are $\sum_1 + \sum_0$ and $\sum_2 + \sum_0$, respectively and they are equal since X is balanced. Hence $\sum_1 = \sum_2$, as required. ■

THEOREM 7.3. *If X is balanced and connected, and if v_r is any node of Γ , then Γ has a spanning tree directed to v_r .*

Proof. The subdigraph consisting solely of v_r satisfies the definition of a tree directed to v_r . So there is a subdigraph of Θ of Γ that is a tree directed to v_r and has the maximum number of nodes consistent with this. Let its node-set be S , with complementary node-set T . Assume T non-null. By 7.2 and the connection of X there is a dart D from T to S . Adjoining D to Θ we obtain a tree larger than Θ directed to v_r . From this contradiction we infer that T is null; Θ is the spanning tree required. ■

We now specialize, first to the case in which the conductances are positive real numbers and later to the case in which each conductance is $+1$, noting that $N(\Delta)$ and $N^*(\Delta)$ satisfy these restrictions.

THEOREM 7.4. *If X is balanced and connected, with positive conductances, then $T(X)$ is positive (by 7.3).*

For all such dinets, Eqs. 6(2) can be solved uniquely by Cramer's rule.

THEOREM 7.5. *Let X be balanced and connected, with positive conductances. Let F be the full flow in X from a source v_a to a sink v_b . Then F has the following properties.*

- (i) v_a has an outgoing dart to a node of lower potential
- (ii) v_b has an outgoing dart to a node of higher potential
- (iii) if any other node v_r has an outgoing dart with a non-zero current, then it has one such dart directed to a node of higher potential and one to a node of lower potential.

Proof. (i) and (ii) follow from 6.2 and 7.4(iii) is a consequence of the first law. ■

COROLLARY 7.5.1. *Let v_r be a node other than v_a and v_b . Then $P(v_r) \leq P(v_a)$ and $P(v_r) \geq P(v_b)$. If either equality holds then the outgoing darts from v_r all carry zero currents.*

Proof. Let S be the set of nodes with the maximal potential P and suppose $P > P_a$. By 7.2 and connection there is a dart outgoing from a vertex v_d of S to a vertex of lower potential. Hence by 7.5(iii) there is one from v_d to a vertex of higher potential, contrary to the definition of S . We eliminate potentials less than P_b analogously. The second part of the corollary now follows from 7.5(iii). ■

We now specialize to unit conductances. In this case we need to make no distinction between X and Γ . We suppose X to be balanced and connected (with unit conductances) and investigate its properties.

We note first that the tree-sum $T(X)$ and the transpedances are all integers, by definition. $T(X)$ is simply the number of spanning trees directed to an arbitrarily chosen node.

Consider the full flow $F = F(a, b)$ in X from source v_a to sink v_b . Its currents are integers. We call their highest common factor the *reduction* $\rho(F)$ of F . It also divides $T(X)$ and $[ab, ab]$, by 6.1 and 6.2. By dividing all the currents of F by $\rho(F)$ we obtain the *reduced flow* from v_a to v_b . Transpedances satisfy the following divisibility rule,

$$T(X) \text{ divides } [bp, bq][br, bs] - [pb, bs][br, bq],$$

where $v_b, v_p, v_q, v_r,$ and v_s are distinct from b .

This rule comes from the expression of $T(X)$ and the transpedances in terms of determinants. If b is the greatest of the five suffixes the four transpedances are cofactors of elements of $K_b(X)$ and the expression on the right is a second order minor of the adjugate matrix of $K_b(X)$. By Jacobi's theorem on the minors of the adjugate it is equal to the product of $\det K_b(X)$ and a suitably signed minor of the transpose of $K_b(X)$. (See, e.g., [1].) The value of this minor being an integer the divisibility rule follows. If b is not the greatest suffix, we may be required to multiply by -1 on the right, but that does not affect the rule.

(Jacobi's theorem is included in the formula for the minors of the inverse matrix of a non-singular square matrix A . For that formula, see, e.g., [5].)

Writing the rule again with t replacing s and subtracting we find, by 6.1, that

$$T(X) \text{ divides } [bp, bq][br, st] - [bp, st][br, bq].$$

Then we can likewise eliminate b from the pair bq , and even from the pair bp , by 6(6). We may thus establish the following.

THEOREM 7.6. $T(X)$ divides $[pq, wx][rs, yz] - [pq, yz][rs, wx]$.

8. S-MAPS

The S-map of this section has many of the properties of an $M_2(\Delta)$ but it is not postulated to have a corresponding S-dissection. We shall find, however, that an S-dissection Δ can be derived from a flow in any S-map M and that M is $M_2(\Delta)$ or a topological equivalent unless the flow has one or more zero currents.

So one way to find a Δ , simplified in the next section, is to draw an S-map and calculate a flow in it.

We define an S-map M as a plane map restricted as follows.

First, we specify the nodes of M as an outer node Ω , some nodes called central and some others called corner-nodes.

Second, the faces of M are to be of three kinds: horizontal, vertical, and sloping. The faces incident with Ω are to be called outsiders.

Third, M is to have the properties listed in the enunciation of 5.1.

Fourth, darts and codarts are to be defined for M as for $M_2(\Delta)$ in Section 5.

THEOREM 8.1. *The S-map M has the properties listed in the enunciation of 5.2.*

Proof. Propositions (ii) and (iii) of 5.2, there deduced from 5.1, follow from our third requirement. Then (i) follows from the fact that each sloping face sharing an edge with a horizontal face F is incident with just one outgoing and just one incoming dart at F and from the dual of this observation. ■

A sloping face is called up-sloping or down-sloping when darts go counter-clockwise or clockwise around it respectively.

For convenience, we suppose M drawn in the closed plane with Ω the point at infinity (See Fig. 6). The four outsiders, taken in clockwise order

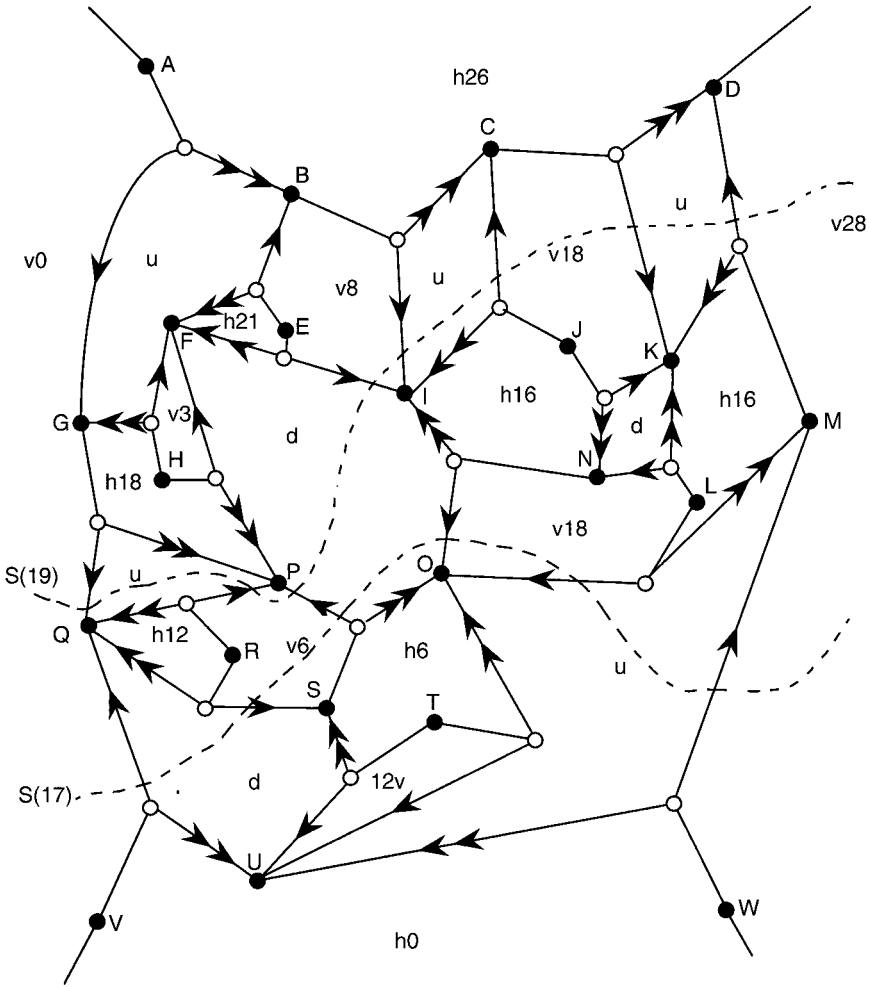


FIG. 6. An S-map.

around the rest of M , are denoted by O_1 , O_2 , O_3 , and O_4 and are referred to as left, upper, right, and lower, respectively. O_1 and O_3 are to be vertical and O_2 and O_4 horizontal. Only at Ω can two vertical or two horizontal faces have a common incident node.

We now define the electrical net $N = N(M)$ and its dual $N^* = N^*(M)$ as in Section 5 for $M_2(\Delta)$. For N and N^* it is convenient to allow the two outsiders, the source and the sink, to have no graphic connection at Ω , despite their geometrical contact there.

THEOREM 8.2. *Let each dart and codart be assigned a unit conductance, then N and N^* are connected balanced dinets to which the theorems of Section 7 apply.*

Proof. N and N^* are balanced by Theorem 8.1. If either were disconnected, the graph of M would be in two parts, each with an edge, meeting only at Ω . This is impossible since the boundary of each face of M is a circuit. ■

By 7.4 we can calculate a flow Q in N from O_2 to O_4 , with a positive current I entering at O_2 and leaving at O_4 . It is usually the full flow or the reduced flow according to theoretical or practical convenience.

From each central node C go a dart D and a codart D^* , constituting a *dart-codart pair*. Each is incident with the sloping face Z at C . D and D^* are incident on their other sides with a vertical face F and a horizontal face F^* , respectively. Q defines a current $i(D)$ in D . We define the cocurrent $i(D^*)$ in D^* as $i(D)$ or $-i(D)$ according as Z is down-sloping or up-sloping.

Let E be one of D and D^* and E^* the other. Write $a(E) = 1$ or -1 when E is dart or codart respectively, $b(E) = 1$ or -1 , when E goes counterclockwise or clockwise around Z , respectively, and $d(E) = 1$ or -1 , when Z is down-sloping or up-sloping, respectively.

Clearly $c(E) = -b(E)$. We can deduce from the preceding definitions and observations that

$$d(E) = a(E) b(E) \quad (1)$$

$$\therefore i(E^*) = a(E) b(E) i(E). \quad (2)$$

Perhaps the most convincing proof is to verify (1) and (2) for each pair of values of $a(E)$ and $b(E)$ separately.

By the *algebraic sum* of the currents (or cocurrents) around any circuit K in a specified direction around K we mean their sum after each current (or cocurrent) has been multiplied by 1 or -1 when its dart (or codart) agrees or disagrees with the specified direction, respectively. So reversing that direction multiplies the algebraic sum by -1 .

THEOREM 8.3. (i) *If the algebraic sum of currents around any sloping face Z is zero, then so is that of cocurrents.*

(ii) *The algebraic sum of cocurrents clockwise around any horizontal face F is equal to the sum of the currents in the darts outgoing from F .*

(iii) *The algebraic sum of currents counterclockwise around any vertical face G is equal to the sum of the cocurrents in the codarts outgoing from G .*

Proof. Proposition (1) follows from the definition of $i(D^*)$ and the fact that the darts all go one way around Z and codarts the other (8.1, as asserting 5.2(iii)).

(ii) and (iii) are consequences of (2). ■

We infer from 8.3 that since the currents satisfy Krichhoff's laws for Q the cocurrents must satisfy those laws for a flow from O_3 to O_1 . Indeed we can assert the following theorem.

THEOREM 8.4. *There is a flow Q^* in N^* from O_3 to O_1 whose currents are the cocurrents of M . The current J entering N^* at O_3 and leaving at O_1 is the (positive) fall of potential from O_2 to O_4 in Q . Likewise, I is the fall of copotential from O_3 to O_1 in Q^* .*

We note that if Q is a reduced flow then so is Q^* .

In Fig. 6 the potential or copotential of each non-sloping face is written within it. Darts are indicated by single arrows and codarts by double ones.

We proceed to the construction of the S-dissection. We take a rectangle R with a horizontal side of length I and a vertical side of length J . We take its lower left corner as the origin of coordinates, with a rightward x -axis and an upward y -axis. x and y are to measure copotential and potential, respectively.

We classify the central nodes as active or inactive according as the incident dart and codart carry non-zero or zero currents and cocurrents (there are two inactive central nodes in Fig. 6). For each active central node C , with its dart D and its co-dart D^* , and with its incident sloping face Z , we define a tile $\tau(C)$, a right-angled isosceles triangle in the plane of R .

At C we have a horizontal face $H(C)$ and a vertical face $V(C)$, incident with D^* and D , respectively. At the other ends of D^* and D we have a vertical face $V_1(C)$ and a horizontal face $H_1(C)$, respectively. Writing P for potential and P^* for copotential we require $\tau(C)$ to have its horizontal side (base) in the line $y = P(H(C))$ and the opposite node (apex) in $y = P(H_1(C))$. Similarly the vertical side (cobase) is to be in $x = P^*(V(C))$ and the opposite node (coapex) in $x = P^*(V_1(C))$. So $\tau(C)$ lies in the square enclosed by these four lines.

THEOREM 8.5. *In the notation of the preceding paragraph $\tau(C)$ has its right corner at $(P^*(V(C)), P(H(C)))$, its apex at $(P^*(V_1(C)), P(H(C)))$, and its coapex at $(P^*(V(C)), P(H_1(C)))$. It is on one side or the other of the line of its hypotenuse according as $i(D)$ is positive or negative. Its hypotenuse is up-sloping or down-sloping in R when the sloping face Z at C is up-sloping or down-sloping in M , respectively.*

Proof. The first sentence is true by construction.

$\tau(C)$ lies on one side or the other of its hypotenuse when it is up-pointing or down-pointing, that is $i(D)$ is negative or positive, respectively.

The hypotenuse of $\tau(C)$ is up-sloping in R if $P(H(C)) > P(H_1(C))$ and $P^*(V(C)) < P^*(V(C_1))$, and if $P(H(C)) < P(H_1(C))$ and $P^*(V(C)) > P^*(V_1(C))$, but not otherwise. Thus the hypotenuse is up-sloping if and only if $i(D) = -i(D^*)$, that is if and only if Z is up-sloping in M . This completes the proof. ■

By 7.5.1. each tile is contained in the rectangle R .

THEOREM 8.6. *Let Z be a sloping face of M . Then the active central nodes of M incident with Z (if any) define tiles with their sloping sides all in the same line $L(Z)$.*

Proof. We enumerate the central nodes incident with Z as C_1, C_2, \dots, C_K in cyclic order around Z in the direction indicated by the darts. Let D_i be the dart and D_i^* the codart from C_1 . Let H_i be the horizontal face incident with D_i^* and V_i the vertical one incident with D_i .

Suppose C_p active. Then by 8.5, $\tau(C_p)$ has its apex at $(P^*(V(C_{p-1})), P(H(C_p)))$ and its coapex at $(P^*(V(C_p)), P(H(C_{p+1})))$. But if C_p is inactive we have

$$P(H(C_p)) = P(H(C_{p+1})), \quad (P^*(V(C_p))) = P^*(V(C_{p-1})).$$

Now let C_q be the next active central node in the dart-direction (supposed distinct from C_p). Whether or not any inactive central nodes intervene between them, it follows from the above observations that the coapex of $\tau(C_p)$ is the apex of $\tau(C_q)$. The hypotenuses of these two tiles, having the same slope by 8.5, lie therefore in a line. ■

As in Section 6 we put $P(O_4) = 0$ and $P^*(O_1) = 0$. Let α be a real number such that

$$0 < \alpha \leq J = P(O_2). \quad (3)$$

Let $A(\alpha)$ be the subdigraph of N made up of the nodes of potential $\geq \alpha$ and of the darts joining them. Similarly let $B(\alpha)$ be defined for nodes of potential $< \alpha$. By 7.5.1 O_2 is in $A(\alpha)$ and O_4 is in $B(\alpha)$.

THEOREM 8.7. *$A(\alpha)$ and $B(\alpha)$ are connected digraphs.*

Proof. Suppose $A(\alpha)$ to have a component K not including O_2 . Let U be the set of nodes of K having the maximum potential in K . Since N is connected and balanced, there is a dart outgoing from some node v of K

to a vertex of $B(\alpha)$, necessarily at a lower potential. Hence, by the first law there is a dart from v to a vertex of higher potential, contrary to the definition of U . The argument for $B(\alpha)$ is analogous. ■

When $A(\alpha)$ and $B(\alpha)$ are considered as parts of M their nodes, of course, are called *horizontal faces*.

A dart or face of M is α -crossing if it has a common incident node with a horizontal face in $A(\alpha)$ and with another in $B(\alpha)$. So an α -crossing face is either vertical or sloping.

We now construct a sequence

$$S(\alpha) = (F_1, D_1, F_2, D_2, \dots, F_{K-1}, D_{K-1}, F_k), \quad (4)$$

alternately of α -crossing faces F_i and α -crossing darts D_j , so that D_i is incident with F_i and F_{i+1} for each $i < k$, these two faces being distinct. So the faces in $S(\alpha)$ are alternately vertical and sloping.

To construct $S(\alpha)$ we put $F_1 = O_1$. Since $J > 0$ we can find an α -crossing dart D_1 incident with F_1 . The sloping face incident with D_1 , necessarily α -crossing, we take as F_2 . On the bounding circuit of F_2 there is at least one other α -crossing dart. We take one as D_2 and the vertical face incident with D_2 as F_3 . We find another α -crossing dart D_3 incident with F_3 and take its incident sloping face as F_4 , and so on until either some face is repeated in $S(\alpha)$ or we arrive at $O_3 = F_k$.

In fact, no face can be repeated. For if one were there would be a set X of α -crossing darts separating one set U of horizontal faces from the complementary set U' of horizontal faces in M . We can suppose O_2 and O_4 to be in U' since there are connected through darts incident with O_3 . But then X must include a dart of $A(\alpha)$ or $B(\alpha)$, by 8.7, and such darts are not α -crossing.

$S(\alpha)$ now extends, without repetition, from $F_1 = O_1$ to $F_k = O_3$. A face F_j is vertical or sloping according as j is odd or even. On one side of $S(\alpha)$, which it seems proper to call either the left or the upper side, we have the horizontal faces of $A(\alpha)$, and on the right or lower side those of $B(\alpha)$, by 8.7. It follows that M has no α -crossing dart or face outside $S(\alpha)$. That sequence is uniquely determined; at each term short of F_k we had only one choice for the next. We call $S(\alpha)$ the α -corridor of M . In 6, $S(17)$ and $S(19)$ are indicated by broken lines traversing them.

We list the numbers that are potentials of horizontal faces, in increasing order, as $\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_t$. Thus $\lambda_0 = 0$ and $\lambda_t = J$.

THEOREM 8.8. *If $\lambda_i < \alpha < \beta < \lambda_{i+1}$, then $S(\alpha)$ and $S(\beta)$ are identical.*

This is because $A(\alpha) = A(\beta)$ and $B(\alpha) = B(\beta)$.

THEOREM 8.9. *Let Z be a sloping face. Let L be an arc in its boundary with vertices enumerated from one end to the other in counterclockwise order around Z as v_1, v_2, \dots, v_q .*

Let H_i and V_i be the horizontal and vertical face, respectively, incident with v_i . Let it be given that all the H_i have the same potential γ and that no dart is directed from v_1 or v_q to a horizontal face of potential $< \gamma$. Then

$$P^*(v_1) \leq P^*(V_q). \quad (5)$$

Moreover, if some dart is directed from v_1 or v_q to a horizontal face of potential $> \gamma$ then the strict inequality holds.

Proof. Suppose there is a dart from v_1 (assumed central). If $D = v_1 v_2$ then $V_1 = V_2$. Otherwise D , incident with V_1 , goes clockwise around Z . Z is thus down-sloping. When the current in D is negative or zero, being one or the other by hypothesis, so is the cocurrent in the codart $v_1 v_2$, by (2). Thus $P^*(V_1) < P^*(V_2)$ in the first case and $P^*(V_1) = P^*(V_2)$ in the second. In both cases $V_2 = V_3$.

Similarly if v_q is central then $V_{q-1} = V_q$ if D is $v_q v_{q-1}$, and otherwise $P^*(V_q)$ is greater than or equal to $P^*(V_{q-1})$ when D carries a negative or zero current, respectively. But now D goes counterclockwise and Z is up-sloping.

If some internal vertex v_r of L is central the current, and therefore the cocurrent, there is zero, by the condition on the H_i . Hence

$$P^*(V_{r-1}) = P^*(V_r) = P^*(V_{r+1}). \quad \blacksquare$$

The theorem follows from these results.

THEOREM 8.10. *If $0 < \alpha < \lambda_1$, then the copotential of the vertical faces is strictly increasing in $S(\alpha)$.*

Proof. Let F_{2j} be a sloping face of $S(\alpha)$. Apart from the darts D_{2j-1} and D_{2j} the bounding circuit of F_{2j} consists of two arcs, W_l and W_r , on the left and right of $S(\alpha)$, respectively. One but not both of the darts is directed from a central end of W_r to a horizontal face of positive potential (being α -crossing). The conditions of 8.9 apply with $L = W_r$, $\lambda = 0$, $V_1 = F_{2j-1}$, and $V_q = F_{2j+1}$. Hence $P^*(F_{2j+1}) > P^*(F_{2j-1})$. \blacksquare

THEOREM 8.11. *Let i, α , and β be such that $\lambda_{i-1} < \alpha < \lambda_i < \beta < \lambda_{i+1}$. Let it be given that copotential is strictly increasing in $S(\alpha)$. Then copotential is strictly increasing in $S(\beta)$.*

Proof. We use the notation on the right of (4) for $S(\beta)$.

Let F_{2j} be a sloping face of $S(\beta)$. Suppose first that it is not α -crossing. Define W_r as in 8.10. Every horizontal face incident with a vertex of W_r lies between $S(\alpha)$ and $S(\beta)$ and so has potential α . One of D_{j-1} and D_j is directed from a central end of W_r .

Hence $P^*(V_{2j-1}) < P^*(V_{2j+1})$, by 8.9.

In the remaining case F_{2j} is α -crossing. If both F_{2j-1} and F_{2j+1} are α -crossing then they are consecutive vertical faces in $S(\alpha)$, and $P^*(V_{2j-1}) < P^*(V_{2j+1})$ by hypothesis.

Suppose next that F_{2j-1} but not F_{2j+1} is α -crossing. Let G be the successor of F_{2j} in $S(\alpha)$ (Fig. 7). Applying 8.9 we find that $P^*(G) \leq P^*(F_{2j+1})$. But $P^*(F_{2j-1}) < P^*(G)$ by hypothesis. Hence $P^*(F_{2j-1}) < P^*(F_{2j+1})$. If F_{2j+1} but not F_{2j-1} is α -crossing (Fig. 7) we argue analogously to the same result.

There remains the subcase in which neither F_{2j-1} nor F_{2j+1} is α -crossing, as in Fig. 7. Then let G be the predecessor and G' the successor of F_{2j} in $S(\alpha)$. Using 8.9 we find $P^*(F_{2j-1}) \leq P^*(G')$ and $P^*(G) \leq P^*(F_{2j+1})$. But $P^*(G') < P^*(G)$ by hypothesis. Hence $P^*(F_{2j-1}) < P^*(F_{2j+1})$. This completes the proof of 8.11. ■

THEOREM 8.12. *If $0 < \alpha < t$ and α is not a λ_i then the copotential of the vertical faces is strictly increasing in $S(\alpha)$, by 8.10 and 8.11.*

THEOREM 8.13. *The tiles corresponding to the active central nodes of M cover the rectangle R without gap or overlap.*

Proof. Choose an α between λ_i and λ_{i+1} , where $0 \leq i < t$. Let R_i be the strip of the rectangle R lying between $y = \lambda_i$ and $y = \lambda_{i+1}$. Consider the α -corridor $S(\alpha)$.

Each dart D_i is incident with an active central vertex C_i .

Making further use of (2) we find that whether C_{2j-1} lies on the right or the left of $S(\alpha)$ the tile $\tau(C_{2j-1})$ has its cobase in $x = P^*(F_{2j-1})$ and its coapex to the right or high-copotential side of that line. But $\tau(C_{2j})$ has its cobase in $x = P^*(F_{2j+1})$ and its coapex to the left of that line. By 8.6 the sloping sides of $\tau(C_{2j-1})$ and $\tau(C_{2j})$ lie on one line.

We infer that the parts within R_i of the tiles corresponding to the darts of $S(\alpha)$ fill that strip without gap or overlap. No other tile overlaps with R_i since no other dart is α -crossing. Since this is true for each R_i and its corresponding α the theorem is established. ■

The e-segment corresponding to a face F can be found by combining the appropriate sides of the tiles $\tau(C)$ such that C is an active central node incident with F . If there is a zero current or cocurrent we may encounter the anomaly of several e-segments (so constructed) meeting at a cross that no

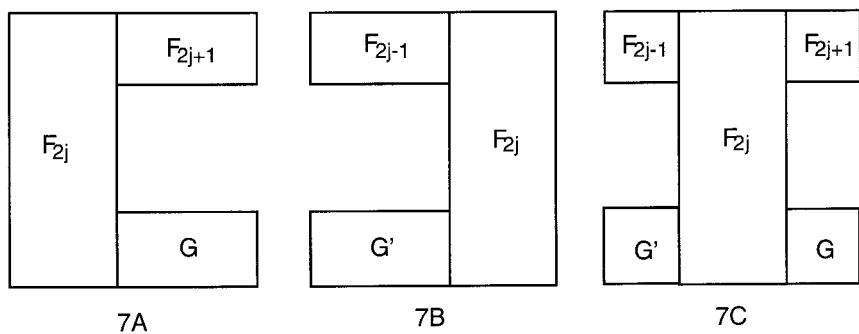


FIG. 7. Relating two corridors.

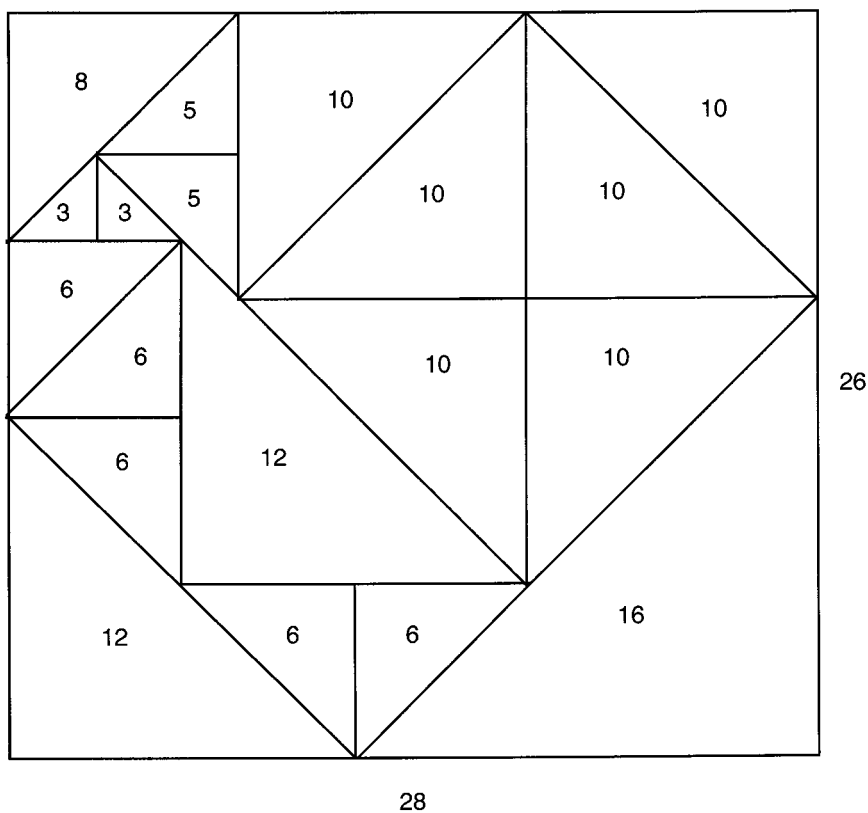


FIG. 8. Rectangle from Fig. 6.

one of them traverses. Figure 8 shows the S-dissection constructed from Fig. 6.

The object of this section is now achieved. But one question obtrudes. What is the relation between the full flows of N and N^* ? The answer is that $T(N) = [O_3 O_1, O_3 O_1]^*$ and that $T(N^*) = [O_2 O_4, O_2 O_4]$, where asterisks refer to N^* . To sketch a proof of the first equation we observe that with Ω -detachment the complementary set of a tree T of N directed to O_2 can be interpreted as a double tree of N^* directed to O_1 and O_3 , the nodes being vertical faces and the edges sloping ones. Each sloping face has just two darts, going opposite ways connecting its two adjacent vertical faces in the tree. Replacing each sloping face by the appropriate one of these darts we get a genuine double tree in N^* , directed to O_1 and O_3 . Similarly given such a double tree in N^* we can work back to a T .

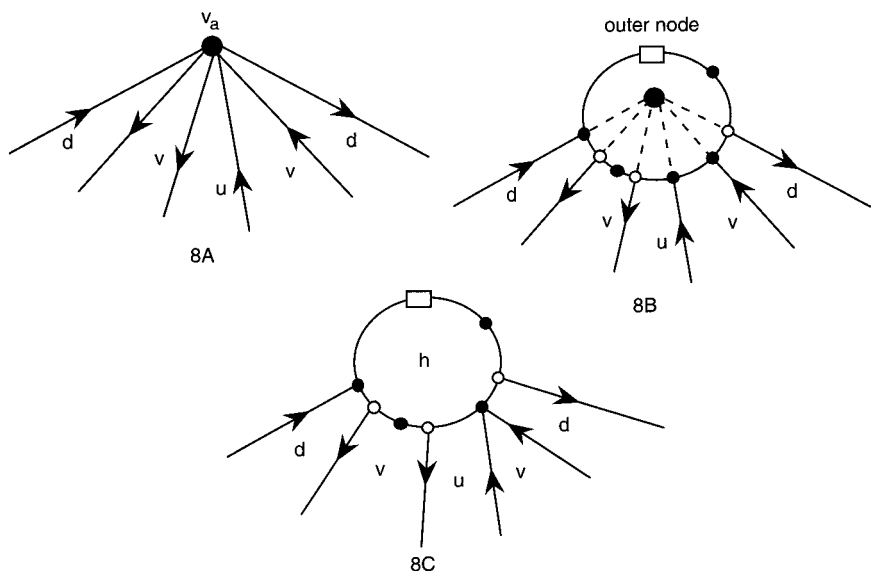
The Δ derived from Fig. 6 is imperfect. It is shown in Fig. 8.

9. CONSTRUCTION OF S-DISSECTIONS

The theory-based method is this. One first constructs an Eulerian plane map N , preferably 2-connected. It is uniquely face-2-colourable. The faces coloured like the outer one are called vertical. The others are sloping. The sloping ones are labelled arbitrarily as up-sloping or down-sloping. The edges are directed clockwise around down-sloping faces and counter-clockwise around up-sloping ones. The edges are now darts.

The Δ -constructor now sees the graph as a connected balanced dinet. He chooses a source v_a and a sink v_b on the boundary of the outer face, solves the Kirchhoff equations, and says that the current in each dart tells the size of a tile in an S-dissection Δ of a rectangle R . Having had some practice, he easily fits those tiles together to make his required Δ .

We, seeking theoretical justification, must pause to expand N into an S-map M so that we may use the theorems of Section 8. To expand a node v_r of N into a horizontal face, we draw a small simple closed curve C around it to cut each incident dart in one point only. The interior of C is to be a horizontal face of M replacing v_r . (See Figs. 9a and 9b.) Where C meets an outgoing dart we place a central node and if, τ being a or b , it crosses the outer face we put an outer node in that stretch. As we go clockwise around C we place a corner-node between each successive ordered pair of central or outer nodes. If C crosses an incoming dart in this stretch we make the corner-node incident with it (or with one of them). If there are two such darts, we slightly distort the second to make it incident with that corner-node (Fig. 9c). There is no third incoming dart there, for by the construction of N any two must border differently sloping faces. Finally,

FIG. 9. Expansion of v_a .

the two outer vertices, arising from v_a and v_b , are pulled together across the outer face to coincide at infinity in a new vertex Ω .

It is easy to verify that the map M so formed satisfies the definition of an S-map with an electrical net equivalent to that of N . The construction of Δ is now justified.

Note that the method gives us no control over the shape of R . For example, only by a fluke will R be a square.

A Δ -constructor will know of an amusing trick. He can reverse all the darts in N to get a new connected and balanced dinet N' . From this he can derive an S-dissection Δ' of a rectangle R' , using the same source and sink as before. Distinguishing quantities referring to N' by primes we observe that the two Kirchhoff matrices are transposes of one another. Using the formulae giving the transpedances of a full flow in terms of determinants and the fact that a matrix has the same determinant as its transpose we infer the following rules.

$$T(N) = T(N'), \quad (1)$$

$$[ab.rs] = [rs.ab]'. \quad (2)$$

Putting $rs = ab$ we discover that R and R' have the same shape. Both have horizontal side $T(N)$ and vertical side $[ab.ab]$. Note that the number of spanning trees directed to a vertex v_a is not altered when all darts are reversed.

10. PERFECT RI-DISSECTIONS

In the Eulerian map of Section 9 no node other than v_a and v_b is allowed to be divalent. This restriction avoids some inconvenient zero currents. An exercise on Euler's polyhedrom formula shows that there must be either a 4-valent vertex, other than v_a or v_b , or a 2-sided face.

If, in the electrical map N of Section 9, there is a digon, then the currents in the digon's two darts are numerically equal and Δ is not perfect. But if the two darts are similarly directed, their two tiles can be combined into a single one with a horizontal hypotenuse. So if we strive for a perfect RI-dissected rectangle, one rule is that no two nodes of N may be joined by more than three darts, and if two or three, then two of them must bound a digon and be similarly directed.

For a tetravalent vertex v , not v_a or v_b , the two outgoing darts carry zero-summing currents. Their two corresponding tiles can be combined into a single one, with a vertical hypotenuse, if and only if the two darts are consecutive in the cyclic order of darts around v . So a second rule is that outgoing darts from a 4-valent vertex must be consecutive there. Moreover, they must not form a 2-circuit; then their zero-summing currents would be equal and so both zero.

Obedience to these rules is found in practice to make a perfectible S-dissection likely.

11. PERFECT DISSECTION OF SQUARES

A perfectly RI-dissected rectangle R can be made into a perfectly RI-dissected square as follows, provided that each of its sides contains sides of at least two tiles. We place two new tiles against two adjacent sides to make a dissected triangle, and then we place a big new tile against the hypotenuse. But such a dissection we dismiss as trivial. We want no main diagonal as a maximal segment and no properly contained RI-dissected rectangle. Such a dissection we call *simple*.

More than 200 simple squares were found empirically by one of us (J.D.S.) by fitting together triangular tiles. Like chess-playing this is a skill to be learned, and when learned it is difficult to explain. The next paragraph attempts a description.

Practice with simple tilings is essential for success with more difficult ones. Lists of tiled polygons with only a few tiles can be made. Often two of them will have the same shape. Then one can be replaced by the other during the construction of a perfect RI-square, usually to remove one of a pair of congruent tiles. Often the construction of a RI-square starts with a

trivial one, split along a main diagonal. Then the tiles on one side, and some others, are removed. The next one tries to triangulate anew the emptied areas. A “greedy” method is recommended. One starts with as big a tile as can be fitted in and seems to simplify the problem (by subjective and unformulated criteria). If that fails the difficult empty polygon is cut into smaller ones that can perhaps be tiled separately. Figure 10 shows a square obtained by this method. It is recorded as 12:14 (J.D.S.). This formula tells us that the reduced side is 12, that the “order” (or number of tiles) is 14, and that the discoverer’s initials are J.D.S.

Recently electrical theory has become relevant to the problem. By reversing all darts in the electrical map of N of a square R we get the electrical map of another square R' , by Section 9. In the notation of that section we have the following theorem.

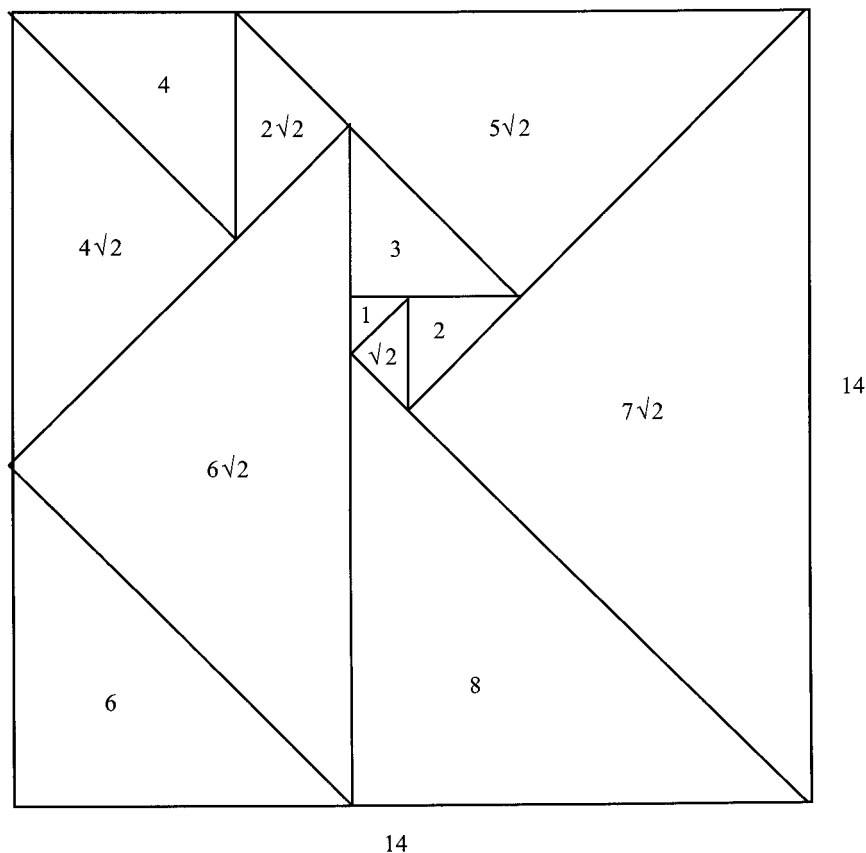


FIG. 10. The square 12:14 (JDS).

of a reduction we can choose distinct nodes v_x and v_y so that p does not divide $[ab, xy]/\rho$. But then

$$p^{\alpha-\beta} \text{ divides } [ab.rs]'$$

for all choices of v_r and v_s by (3). Thus $p^{\alpha-\beta}$ divides ρ' by the definition of a reduction. Accordingly

$$p^\alpha \text{ divides } \rho\rho'.$$

Since this holds for every prime factor p of $T(N)$ the theorem is proved. ■

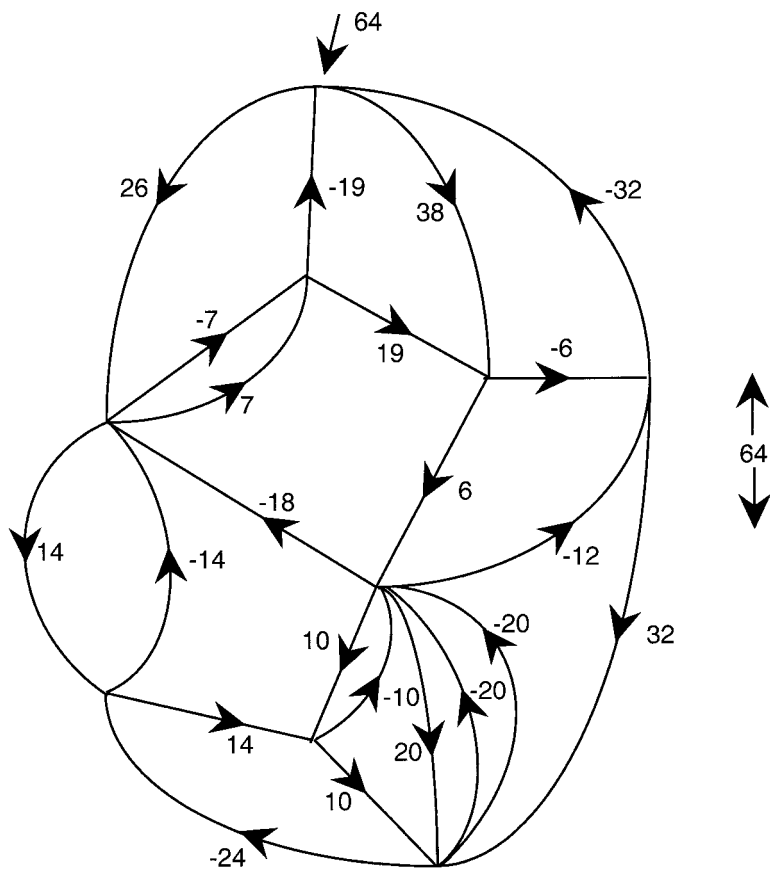


FIG. 12. From Fig. 11 by dart-reversal.

When the S-dissection of R is perfectible (by repeatedly uniting two congruent tiles to make a bigger tile) one expects ρ to be small. If so, 11.1 indicates that ρ' must be large. This seems to explain the disappointing results of attempting to get a new perfect square from an old one. But “you must always invert.” “Collapsed” S-dissections of squares, those with small reduced horizontal sides, are likely to have crosses and excessively many imperfections, but they are easier to find than perfectible ones. Given a collapsed S-dissection of a square perhaps we can get a perfectible one from it by reversing darts.

For someone attempting this we recommend the following procedure. First the obvious electrical map is drawn. Then one or more nodes are split, each into two or more new nodes in such a way that Kirchhoff's laws are still satisfied with the original currents. The resulting net is not required

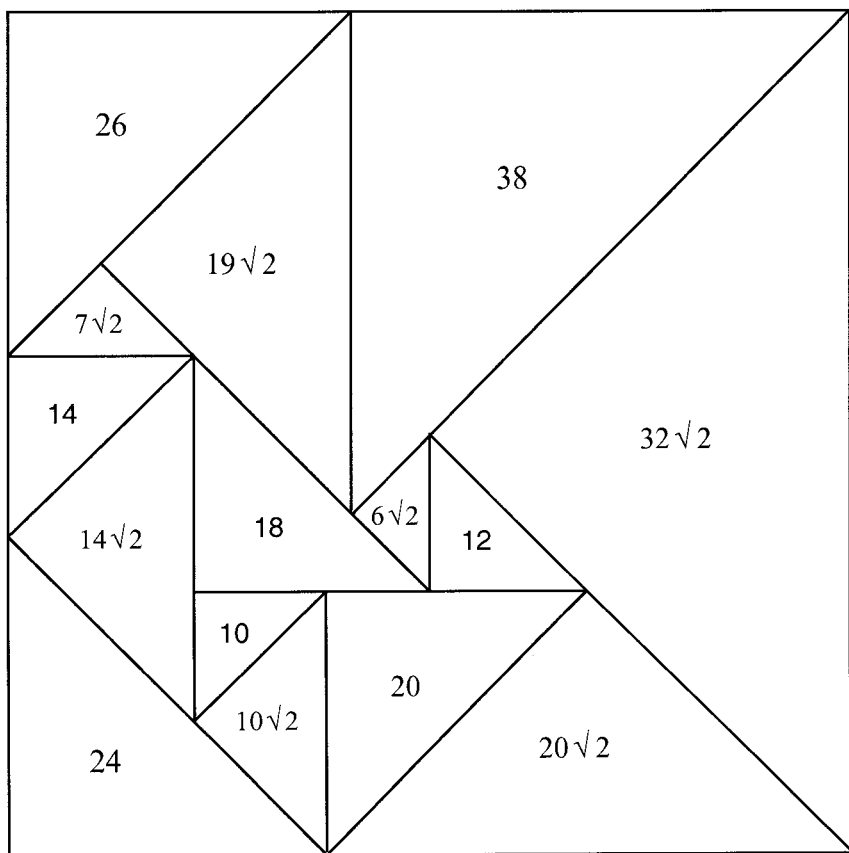


FIG. 13. Perfect square from Fig. 12. 15:64 (WTT).

to be balanced. Next, in one or more operations we join two nodes of equal potential by a dart with zero current or we join three or more equipotential nodes by means of a more complicated structure. The final result must be balanced. Figure 11 shows a collapsed square and an electrical map derived for it by the method just described. The next step is to reverse all the darts and derive the corresponding S-dissected square. The tricky part is ensuring that the new dinet satisfies the perfectibility rules of Section 10.

Figure 12 shows the reduced flow in the new dinet and Fig. 13 shows the corresponding RI-dissected perfect square.

12. CONNECTIONS WITH EARLIER WORK

The squared rectangles of [2] are described by electrical maps with undirected edges. They become S-dissections when each constituent square is diagonally bisected into two congruent triangles. Then each edge in an electrical map is replaced by two oppositely directed darts bounding a digon, thus giving a map identifiable with the N of Section 9. The reverse-dart method gives nothing new, merely switching diagonal bisectors in each square.

The dissections into equilateral triangles of [3] and [6] can be regarded as sheared S-dissections in which each sloping e-segment is up-sloping. In a corresponding N each dart goes counter-clockwise around the incident sloping face and clockwise around the vertical one. It follows that M is *alternating*, incoming and outgoing darts alternate at each vertex. We are dealing with the special case of alternating electrical maps.

The rule of perfection is relaxed in this case. Two congruent tiles are admissible provided that their currents differ in sign. Usually two such tiles

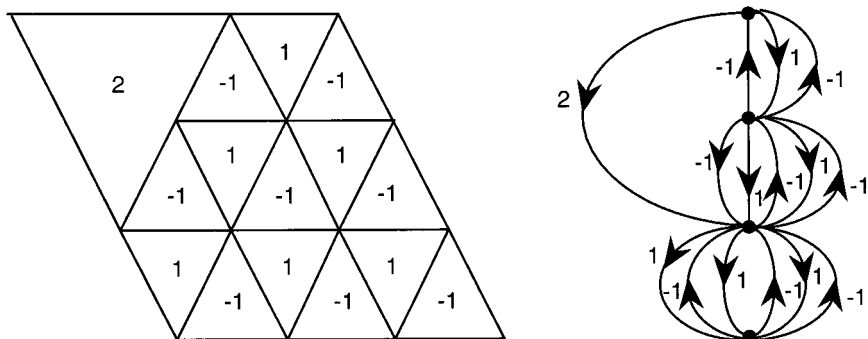


FIG. 14. A collapsed rhombus and its electrical map.

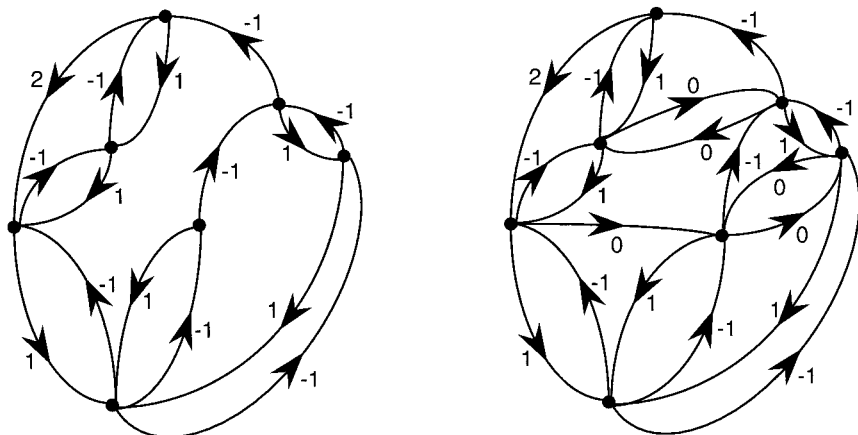


FIG. 15. Modifications from Fig. 14.

(as equilateral triangles) fit together to make a rhombus, or sheared square. So a perfect rhombus can be got by shearing a perfect squared square.

The author of [6] hoped for a perfect rhombus not of that kind, not the trivial one of two congruent tiles, nor yet of the "trivial" kind got by adding three big new tiles to a perfect parallelogram. Only recently did he notice that there is a simple method of finding one. Figure 14 shows a collapsed rhombus with its obvious dinet. Figure 15 shows modified dinets of the rhombus, obtained by operations of the kind recommended in Section 11. Figure 16 shows the dinet got by reversing all darts, with its reduced flow, and Fig. 17 shows the corresponding perfect rhombus.

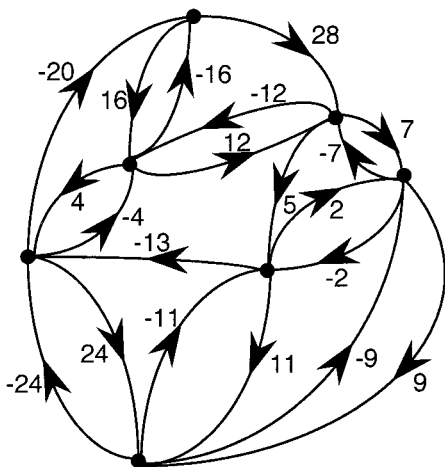


FIG. 16. Derived from Fig. 15 by reversing darts.

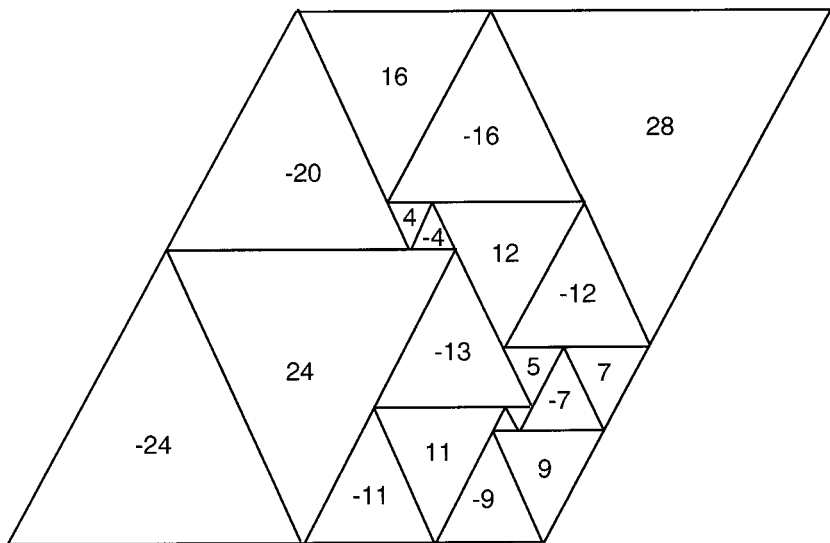


FIG. 17. Perfect rhombus from Fig. 16.

13. STATE OF THE ART

Many new perfect RI-squares have been found by the method of reversing darts. One of them, shown in Fig. 18, has only 11 tiles.

It is not necessary to make a drawing of every dissected square or rectangle deemed worthy of mention. A formula analogous to that introduced by C. J. Bouwkamp for describing rectangles dissected into squares can be used. Thus to describe the S-dissection of Fig. 1B we can write

$$(-3, 3, 3, -5)(1, -1, 2)(-6, 1)(5). \quad (1)$$

Here the first bracketed expression lists the currents or signed sizes of the tiles incident with the upper side of R , in the order of the tiles from left to right. Positive numbers indicate side-incidence and negative ones corner incidence.

The second pair of brackets refers to the next horizontal e-segment down and to the sequence of signed sizes of the tiles incident with it and situated below it. The expression $(-6, 1)$ refers likewise to the third horizontal e-segment down and (5) likewise to the fourth.

The rule is to take the horizontal e-segments in order from top to (just short of) bottom of R . Should two or more be on the same level they are to be taken in order from left to right.

From (1) we can reconstruct the S-dissection, tile by tile in the order of writing. Whether a tile is left-pointing or right-pointing is determined by the necessity of fitting against one or more of its predecessors. Except of course for the first tile, but that is necessarily right-pointing.

To describe the RI-dissection of Fig. 1A we modify the formula as follows.

$$(-3, 3^*, 3^*, -5^*)(1, -1^*, 2)(-6, 1^*)(5) \quad (2)$$

Here an asterisk indicates a tile that is to be united with a neighbour of the same absolute size.

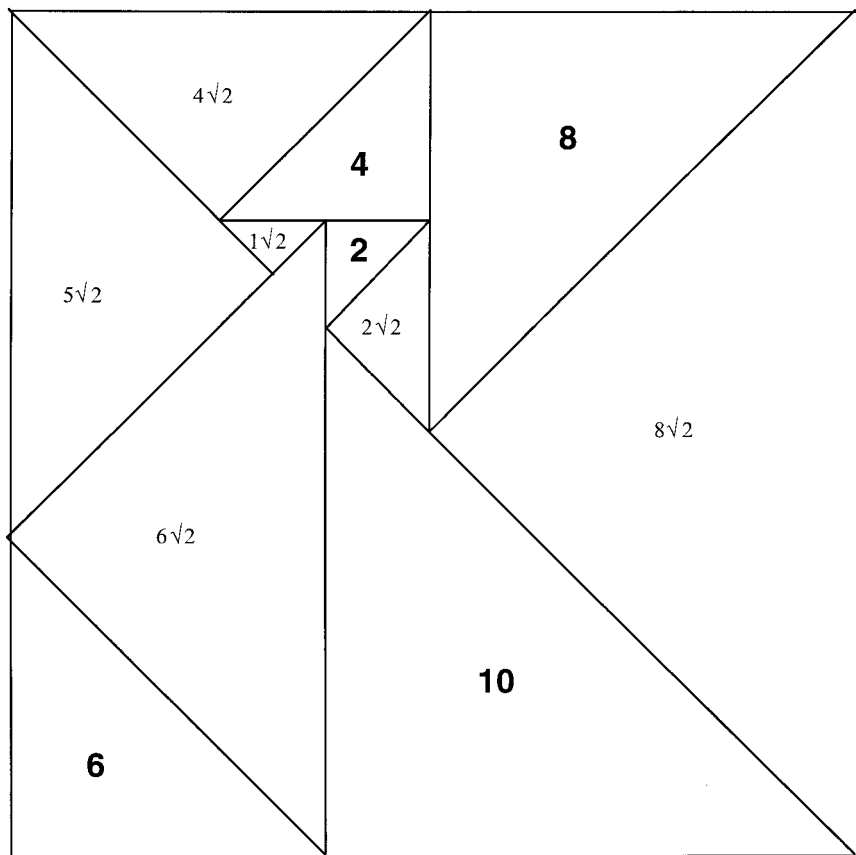


FIG. 18. 11:16 (JDS).

As another example we give the “Bouwkamp formula” for the square 11:16 (JDS). (See Fig. 18.)

$$(-5^*, 4^*, 4^*, -4, 8, -8^*)$$

$$(1^*, 1^*, -6^*, 2, -2^*)(5^*)(-10, 2^*)(8^*)(-6, 6^*) \quad (3)$$

For dissections given as perfectible it may be thought sufficient, as in the following table, to give an asterisk to only one of two tiles to be united or if there are only two tiles of the same absolute size to give them no asterisk at all.

We give below the descriptions of some other perfect squares. Some of these are *congeners*, that is, squares with equal reduced sides or with their reduced sides in a simple numerical ratio. Congeners are got by the method of reversing darts, with slightly different ways of expanding the “trivial” electrical net.

SOME LOW ORDER SIMPLE PERFECT RIGHT-ANGLED ISOSCELES TRIANGLED SQUARES

11: 16A (JDS) $[-5^*, 4^*, 4, 4, 8, -8^*] (1^*, 1, -6^*, 2, -2^*) (5) (-10, 2) (8) (-6, 6)$

12: 14A (JDS) $(-4^*, 4, -2^*, 5^*, 5, -7^*) (2, -6^*, 3) (4) (1, -1^*, 2) (-8, 1) (7) (-6, 6)$

14: 24A (JDS) $(-7, 7^*, 7, -5, 10, -12^*) (-2, -3, 5) (9, -1) (-8^*, -16, 4, -2^*) (2) (12) (-8, 8)$

14: 28A (JDS) $(-9^*, 9, 5, -5^*, 14, -14^*) (-3, 5) (1, -10^*, 2, -2^*, -2) (9) (-18, 4) (14) (-10, 10)$

14: 32A (JDS) $(-10^*, 10, -5^*, 6, 16, -16^*) (5, -1^*, -1) (-2^*, 2, 2) (-12^*, 4, -4^*) (10) (-20, 4) (16) (-12, 12)$

14: 40A (JDS) $(-13^*, 10^*, 10, -10, 20, -20^*) (3^*, 3, -2^*, 4, -5^*) (-14^*, -2, 2) (13) (-26, 1^*, 1) (5) (20) (-14, 14)$

14: 40B (JDS) $(-13^*, 13, 7, -7^*, 20, -20^*) (-3^*, 7) (3, -2^*, 4) (-14^*, -2, 2) (13) (-26, 6) (20) (-14, 14)$

14: 40C (JDS) $(-10, 10^*, 10, -7^*, 20, -20^*) (-3^*, -3, 7) (16, -2^*, 4) (-14^*, -2, 2) (-26, 6) (20) (-14, 14)$

14: 40D (JDS) $(-13^*, 10^*, 10, -8, 20, 20^*) (-2^*, -4, 4^*, 4, -6^*) (2) (1^*, 1, -14^*, 2) (13) (-26, 6) (20) (-14, 14)$

14: 40E (JDS) $(-13^*, 12^*, 12, -8^*, 16, -20^*) (-4, -8, 8) (1^*, 1, -14^*, 2, -2^*) (13) (-26, 2) (4^*, 4) (20) (-14, 14)$

15: 24A (JDS) (4, -4^* , 4, 10^* , 10, -10) (8, -8^* , -8) (2, -1^* , 1, 7, -7^*) (2, -3^*) (12, -12^* , -12 , 2^* , 2) (3) (7)

15: 36A (JDS) (-11 , 11^* , 11, -14^* , -14 , 14) (5, -5^* , -5 , 3^* , 3) (2, 4, -4^* , -4 , 22) (-10^* , 10) (8, -18^* , -18) (10)

15: 38A (JDS) (-12^* , 12, -6^* , 13^* , 13, -19^*) (6, -4^* , -7) (-14^* , -4 , 4) (12) (1, -3^* , 6) (-24 , 2^* , 2) (3) (19) (-14 , 14)

15: 40A (JDS) (-13^* , 13, 7, -7^* , 20, -20^*) (-4 , 7) (-1^* , 1, 3, -3^* , 3) (1, -1 , 4^*) (13) (-26 , 6) (20) (-14 , 14)

15: 44A (WTT) (18, -13^* , 26, -30) (-5^* , 5, 13) (10, -10^* , -8) (-12 , 4^* , 4) (-16 , 10) (8^* , 8, -14^* , -14 , -14) (6^* , 6)

15: 44B (JDS) (-12 , 11^* , 11, -10 , 22, -22^*) (-1^* , -2 , 2^* , 2, -9^* , 6, -6^*) (1) (7^* , 7) (-28 , 6) (9) (22) (16, -16)

15: 44C (JDS) (18, -13^* , 26, -22^*) (-5^* , -5 , 13) (10, -10^* , 8) (-8^* , -22 , 22) (-12 , 4^* , 4) (-16 , 10) (8) (6^* , 6, -6)

15: 48A (JDS) (8, -8^* , -8 , 20^* , 20, -20) (16, -16^* , -16) (3^* , 3, -2^* , 14, -14^*) (-1^* , -2 , 2) (1) (24, -24^* , -24 , 10) (14)

15: 64A (JDS) (26, -19^* , 38, 32^*) (-7^* , -7 , 19) (14, -14^* , -18) (-6^* , -12 , 32) (6) (-24 , 14) (10, -10^* , 20, -20^* , -20) (10)

15: 72A (JDS) (40^* , 14, -9^* , 18, -18^*) (-5^* , -5 , 9) (10, -22^* , -22) (18) (-16^* , -16 , 12^* , 12) (4^* , 4, -36^* , -36 , 36) (32)

15: 76A (JDS) (-24^* , 24, -12^* , 26^* , 26, -38^*) (-12 , -8^* , -14) (-28^* , -8 , 8) (24) (-2 , 7^* , 7) (-48 , 10, -5) (5) (38) (-28 , 28)

RI-dissections of cylinders, preferably with height equal to circumference, have also been studied. One is shown in Fig. 19. The general S-dissected cylinder has an electrical net like that of Fig. 4, but with source and sink separated by some circuit. It can be given a Bouwkamp formula, with the first bracketed expression going all round the cylinder. It is best to start

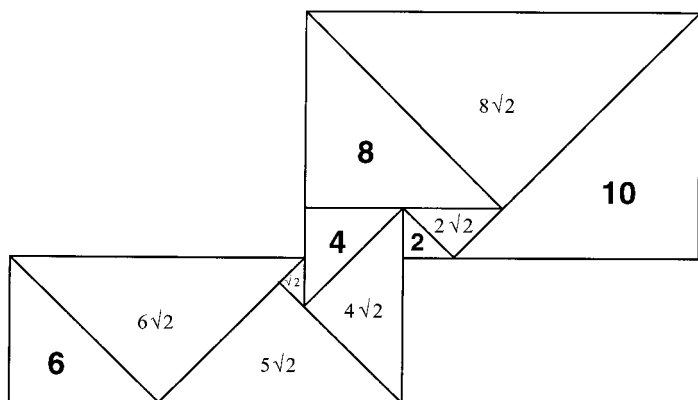


FIG. 19. 11:16 Cylinder (JDS) height = circumference.

and end that sequence at some vertical e-segment. The description of Fig. 19 is as follows:

$$(-8, 8^*, 8^*, -10)(4, -4^*, -2, 2^*, 2^*)(-6, 6^*, 6^*, -1^*)(-5^*, -5^*, 1^*)(4^*). \quad (4)$$

We have found some infinite sequences of perfect RI-squares. From these, and two squares above, we deduce that there is a simple perfect square of every order ≥ 11 . (“Simple” means with no properly contained dissected rectangle or triangle.)

REFERENCES

1. A. C. Aitken, “Determinants and Matrices,” Edinburgh, 1939.
2. R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, The dissection of rectangles into squares, *Duke Math. J.* **7** (1940), 312–340.
3. R. L. Brooks, C. A. B. Smith, A. H. Stone, and W. T. Tutte, Leaky electricity and triangulated triangles, *Philips Res. Rep.* **30** (1975), 204–219.
4. S. Chaiken, A combinatorial proof of the All-Minors Matrix-Tree Theorem, *SIAM J. Algebra Discrete Meth.* **3** (1982), 319–329.
5. R. A. Horn and C. A. Johnson, “Matrix Analysis,” Cambridge University Press, Cambridge, UK, 1985.
6. W. T. Tutte, The dissection of equilateral triangles into equilateral triangles, *Proc. Cambridge Philos. Soc.* **44** (1948), 463–482.
7. W. T. Tutte, “Graph Theory as I have known it,” Oxford, 1998.