# Aperiodic Tiles*

Robert Ammann,[1] Branko Grünbaum,[2] and G. C. Shephard[3]

[1] P.O. Box 265, Billerica, MA 01821, USA

[2] University of Washington GN-50,
Seattle, WA 98195, USA
grunbaum@math.washington.edn

[3] University of East Anglia, Norwich,
Norfolk NR4 7TJ, England

**Abstract.** A set of tiles (closed topological disks) is called *aperiodic* if there exist tilings of the plane by tiles congruent to those in the set, but no such tiling has any translational symmetry. Several aperiodic sets have been discussed in the literature. We consider a number of aperiodic sets which were briefly described in the recent book *Tilings and Patterns*, but for which no proofs of their aperiodic character were given. These proofs are presented here in detail, using a technique with goes back to R. M. Robinson and Roger Penrose.

## 1. Introduction

Since their discovery in the 1970s, the tiles known as Penrose's kites and darts have become well known and there already exists a considerable literature about them (see [6] for details and references). The reason for the widespread interest in these tiles is not surprising; they have the remarkable property of being *aperiodic*, and this property is important in fields ranging from recreational mathematics [5] to solid-state physics [9], [10]. Moreover, it leads to a variety of intriguing mathematical problems.

To explain the term *aperiodic* we briefly review the necessary definitions. By a *tiling* we mean a family $\mathbf{T} = \{T_1, T_2, \ldots\}$ of sets (the *tiles*) which cover the plane without gaps or overlaps. If all the tiles are congruent (directly or oppositely) to one of a minimal set $\mathbf{P} = \{P_1, P_2, \ldots, P_n\}$ of tiles, then these are called the *prototiles*

of the tiling **T**, and we say that they *admit* the tiling **T**. The *symmetries* of **T** are the isometries of the plane that map **T** onto itself, and we say that **T** is *nonperiodic* if it does not have any translation as a symmetry. A set **P** of prototiles is called *aperiodic* if every tiling that **P** admits is nonperiodic.

It is not so well known that about 2 years after Penrose's discovery of his aperiodic set (the kite and the dart), one of us (R.A.) found several aperiodic sets of prototiles, which he communicated to the other two authors. These sets are described on pp. 550–558 of [6], but without any proofs concerning their aperiodic character. The main purpose of this paper is to establish that these sets of prototiles are actually aperiodic. We follow the terminology and notation of [6], and discuss what are called there *Ammann sets* 2, 3, 4, and 5. (The Ammann set number 1 [6, p. 529] is of a rather different nature and does not concern us here.) In order to make this paper self-contained we define and describe these sets, denoted by **A2**, **A3**, **A4**, and **A5**, and we shall prove:

**Theorem.**   *Each of the sets of prototiles* **A2, A3, A4,** *and* **A5** *is an aperiodic set.*

We devote one section of the paper to each of the sets. In Section 2 we consider the set **A2** giving full details of the proof. In subsequent sections we outline the methods only, leaving the verification of certain parts of the argument for the reader to fill in. The final section is devoted to general remarks and observations.

Another major aim of this paper is to provide some nontrivial examples for the use of the "composition technique" for establishing that a set of prototiles is aperiodic. This technique, to be explained presently, was originally used by Penrose [12] and Robinson [14] to prove the aperiodicity of the sets of prototiles they discovered.

To begin with we must explain, in general terms, how we shall proceed. The proofs depend essentially on the process of *composition*, which we describe here for the set **A2**. This set contains two tiles, as shown in Fig. 1(b). The markings on the tiles are parts of ellipses, and the *matching condition* is that the tiles must be put together in such a way that complete ellipses are formed; an example of part of a tiling **T** by these prototiles is given in Fig. 3. In Fig. 2 we show two *patches* of tiles, one containing two tiles and the other containing three. If we think of the tiles in each of these patches as being welded together we obtain two larger tiles (*supertiles*) which are similar in shape to the prototiles from which we started. In Fig. 3 heavier lines indicate how the tiles in **T** may be grouped into patches each of which is like one shown in Fig. 2. Hence, after welding the tiles together, **T** is transformed into a tiling by the corresponding supertiles. Our proof that a given set **P** of prototiles is aperiodic depends essentially on this process of composition, that is, on the transformation of a tiling into one by supertiles. To be precise, for any **P** we must show:

  (a) in every tiling admitted by **P** there is a *unique way* in which the tiles can be grouped into patches which lead to a tiling by supertiles as explained above; and

  (b) the markings on the supertiles, inherited from the original prototiles, imply a matching condition for the supertiles which is exactly equivalent to that originally specified for the prototiles.
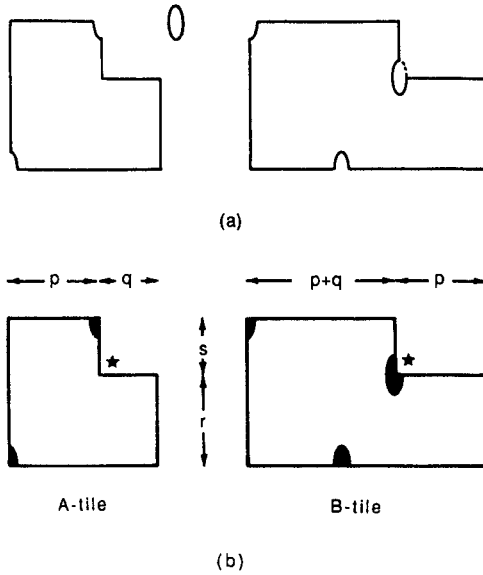
(a)

(b)

**Fig. 1**

We conclude this introductory section with a short explanation of why conditions (a) and (b) imply aperiodicity. Suppose **P** admitted a tiling **T** which was periodic, say it admitted as a symmetry a translation through a distance $L$ in some direction. Then by composing **T** we would obtain a tiling by supertiles which would also admit as a symmetry the same translation through distance $L$; for this step the uniqueness assumption (a) is essential. Because of (b) the composition process may be repeated as often as we wish, and so we can obtain, from **T**, tilings by larger and larger tiles, each with this same translational symmetry. Clearly, this is a contradiction, since if the incircles of all the tiles in a tiling have diameters greater than $L$, then it is impossible for a translation through distance $L$ to be a symmetry of that tiling. This contradiction shows that our original assumption that **T** has a translational symmetry must be false, and since this is true for every tiling admitted by **P**, it follows that **P** is an aperiodic set.



(a)                                                                        (b)

**Fig. 2**

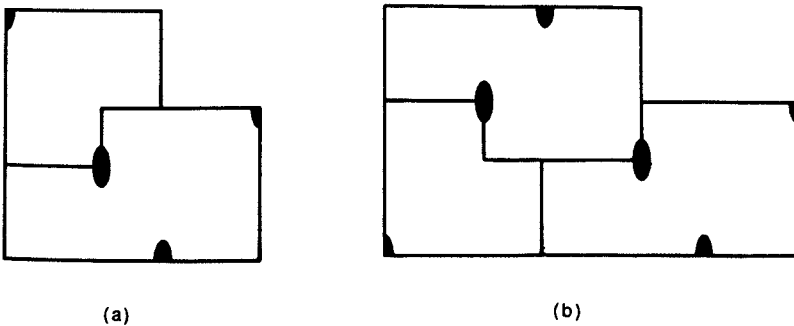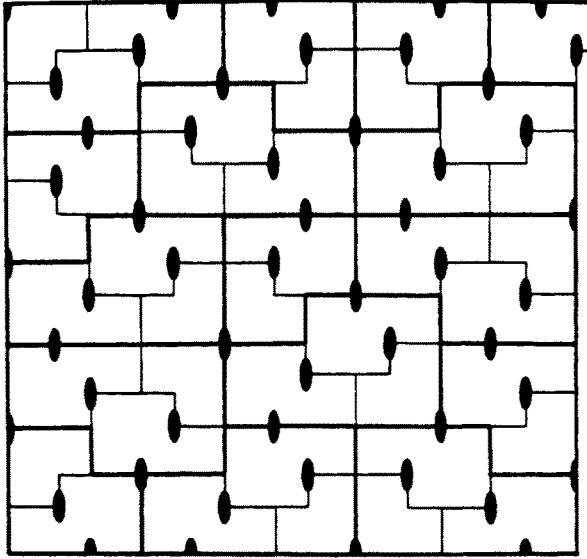**Fig. 3**

Finally, a word of practical advice to the reader. Since a major part of the following proofs consists in verifying that tiles can, or cannot, fit together in certain specific ways, we strongly suggest that the reader makes copies of each of the tiles (about half a dozen of each), preferably on some kind of transparent material so that they can be turned over without hiding the markings. Then it is possible to verify, quickly and without difficulty, all our assertions even when lack of space prevents us from giving details of justification in the text.

## 2. The Aperiodic Set A2

This set can be defined in several equivalent ways, two of which are shown in Fig. 1. In Fig. 1(a) we show a set of three (unmarked) tiles, one of which is in the shape of a small ellipse and may be called a *key tile*, in that its sole purpose is to ensure that the other two tiles fit together in the correct manner. The choice of an ellipse is arbitrary—any other tile which is sufficiently small, has symmetry group $d2$, and does not tile by itself could be used instead. In Fig. 1(b) we show an equivalent set of two prototiles with *markings* in the form of parts of ellipses. As mentioned in the Introduction, the matching condition is that the tiles must be placed together so that complete elliptical markings are formed (see Fig. 3). The second variant is more convenient from our point of view and is used in the following discussion. The prototiles are referred to as the A-tile and the B-tile.

The dimensions of the tiles are indicated in Fig. 1(b), where $p$, $q$, $r$, and $s$ are arbitrary positive lengths. It is convenient to specify that $p/q = r/s = \tau$ (the *golden number* $\tau = (1 + \sqrt{5})/2 = 1.618034...$) because then the supertiles, built up from the prototiles as shown in Fig. 2, are similar (in the sense of Euclidean geometry)

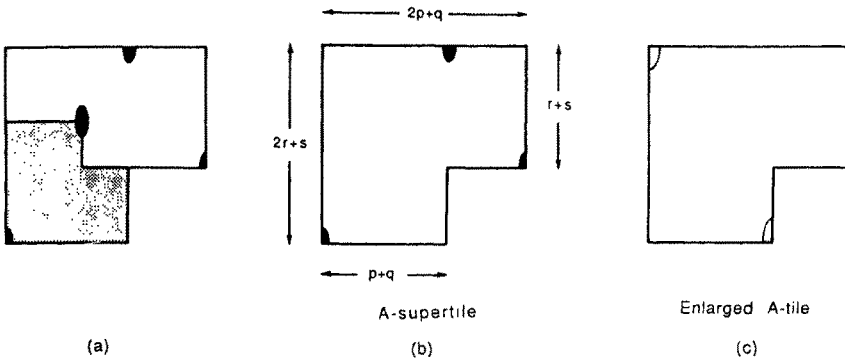A-supertile    Enlarged A-tile

(a)    (b)    (c)

**Fig. 4**

to the prototiles—in fact, they are exactly $\tau$ times as large. However, this is a matter of convenience only; the same arguments hold, with obvious modifications, whatever the values of $p$, $q$, $r$, and $s$. It is worth remarking that whilst there is, *a priori*, no restriction on the orientation of the tiles, the elliptical markings and the matching condition ensure that only reflections in horizontal and vertical mirrors and rotations through 180° are permitted. Rotations through 90° can never be used.

We begin by showing that any tiling **T** by these prototiles is built up uniquely from the patches shown in Fig. 2. Consider which tiles can fit into the reflex angle of the A-tile marked by a star in Fig. 1(b). There are very few possibilities: an A-tile has three unmarked corners and a B-tile has four, and it might be thought that any of these could fit into the reflex angle. However, this is not so since in all cases except one, the matching condition specified by the quarter ellipse on the upper arm of the original A-tile is violated. The only possible arrangement of tiles is shown in Fig. 4(a). Since *every* A-tile must be associated with a B-tile in this way, we may think of them as welded together so as to form the tile shown in Fig. 4(b). In view of Fig. 2(a) we shall call this an A-supertile. Thus we have converted any tiling by A-tiles and B-tiles into (a uniquely determined) one by A-supertiles and B-tiles.

Now we consider which tiles can fit into the reflex angle of the B-tile (marked by a star in Fig. 1(b)). For this we need a tile a corner of which is marked with a quarter ellipse, so it appears that there are three possibilities, namely those shown in Figs. 5(a)–(c); either we use another B-tile, or we use an A-supertile in one of two positions. However, it is easily verified that, in the case of the arrangments shown in Figs. 5(a) and (b), the tiling cannot be continued; that is to say, it is impossible to fit further tiles round the given pair of tiles without leaving gaps. To be precise, for the arrangment in Fig. 5(a) whatever tile is fitted into the reflex angle marked ○ (and there are, of course, three possibilities, one with a B-tile and two with A-supertiles) there is no way in which another tile can be fitted into the position marked ●. (Notice that both reflex angles can be filled simultaneously by an A-tile, but this is not relevant since we recall that only tilings by B-tiles and *A-supertiles* are under consideration.)
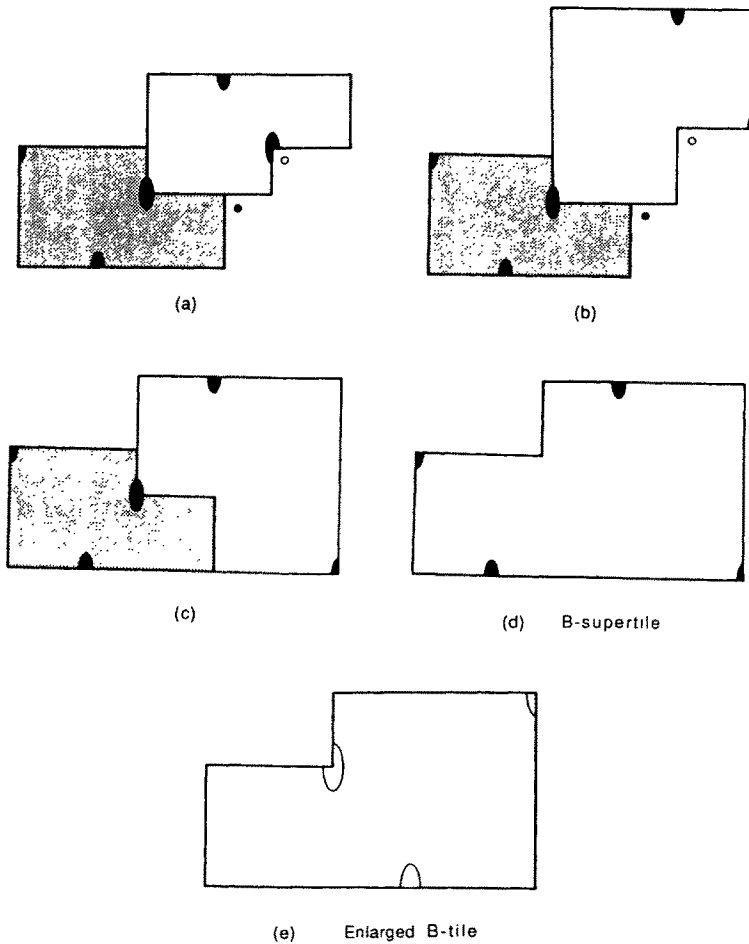
(a)



(b)



(c)



(d)    B-supertile



(e)     Enlarged B-tile

**Fig. 5**

Similar arguments apply to the arrangement shown in Fig. 5(b). Here again there are three possibilities for filling the reflex angle o, two using a B-tile and one using an A-supertile, but in no case can a tile be found to fill the position •, so the tiling cannot be continued. We deduce that every B-tile must be associated with an A-supertile as in Fig. 5(c). Hence we weld these two together and so obtain the tile of Fig. 5(d) which, because of Fig. 2(b), we call a B-supertile. We have thus completed the first part of our proof by showing that any tiling by the set A2 can be uniquely transformed by composition into a tiling by supertiles.

For the second part of the proof we observe that if the supertiles were marked as in Figs. 4(c) and 5(e), then the matching condition for the supertiles would be exactly the same as for the A-tiles and B-tiles (since these markings are simply enlargements of those on the original tiles). We must therefore show that the matching condition inherited from the A-tiles and B-tiles, shown in Figs. 4(b) and
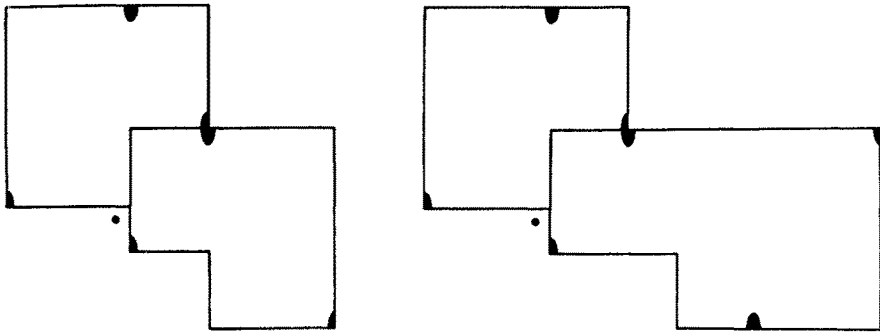
Fig. 6

5(d), is exactly equivalent to this. There is no difficulty in verifying that every arrangement of tiles allowed by the markings in Figs. 4(c) and 5(e) is also permitted by the markings of Figs. 4(b) and 5(d), but the converse statement is not immediate. Several additional possibilities arise such as those shown in Fig. 6. However, we can readily verify that in each case of the tiling cannot be continued, and so these patches cannot occur. (To be precise, in each of the cases shown no tile will fit into the reflex angle marked ●.)

This concludes the proof of the aperiodicity of the set **A2** of prototiles.

## 3. The Aperiodic Set A4

The tiles of this set are shown, in two variants, in Fig. 7. In Fig. 7(a) there are three tiles, one of which (the key tile) is in the shape of an arrow; its only purpose is to ensure that the other tiles fit together in the correct manner. In Fig. 7(b) we show two tiles with markings that specify the matching condition; the markings must fit together to form arrowshapes like the key tile of Fig. 7(a). This second set is more convenient for our purpose, and is used in the following discussion.
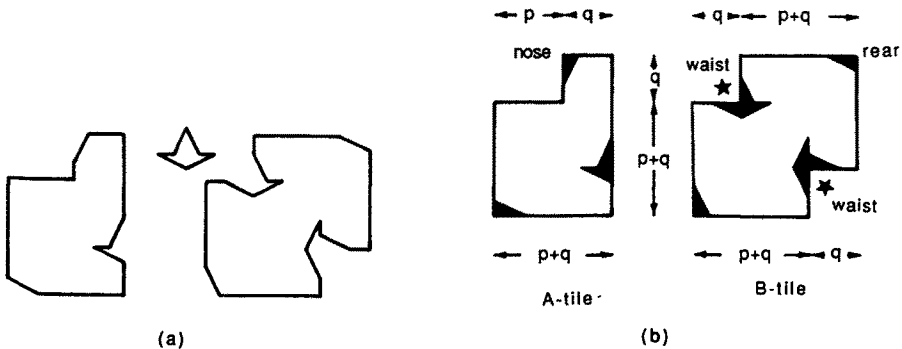


(a)
(b)

Fig. 7

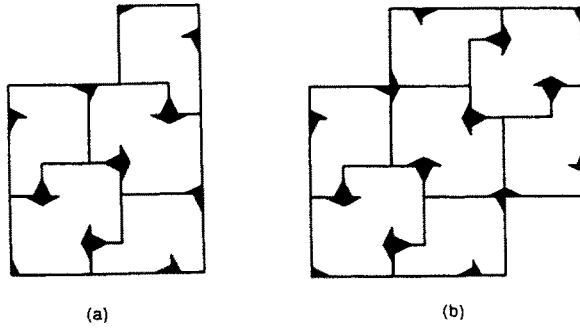(a)                                        (b)

**Fig. 8**

In Fig. 7(b) the dimensions of the tiles are shown. Here $p$ and $q$ may take arbitrary positive values, but if $p/q = \sqrt{2}$, then the supertiles, built up from the prototiles as in Fig. 8, are similar to the prototiles—in fact, they are exactly $1 + \sqrt{2}$ times as large. The same arguments leading to the proof of aperiodicity apply with obvious modifications for all values of $p$ and $q$, but then the word "similar" has to be interpreted appropriately. In Fig. 9 we show an example of a tiling by this set of tiles, with the patches of tiles corresponding to the supertiles marked by heavier lines.

To start the proof we consider tiles which will fit into either "waist" of the B-tile, marked by stars in Fig. 7(b). Clearly, we must use the corner of a tile that is marked with a quarter arrow, and it is immediate that there are four possibilities shown in Figs. 10(a)–(d). However, the two shown in Figs. 10(b) and (d) cannot be continued since no tile will fit into the reflex angles marked ●. This applies to each of the two waists of a B-tile, so we deduce that there are exactly four



**Fig. 9**

(a)                                   (b)
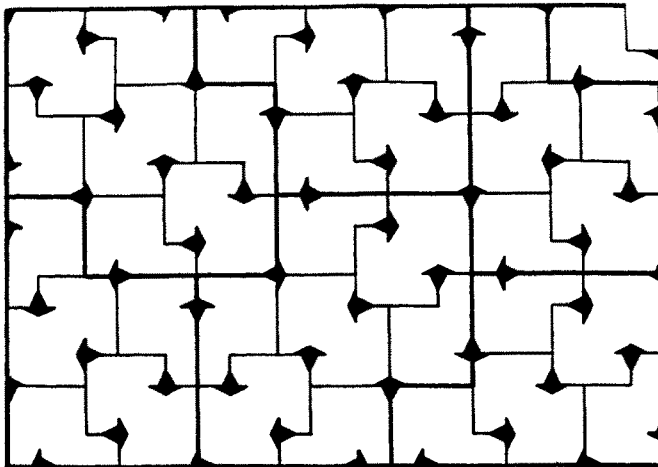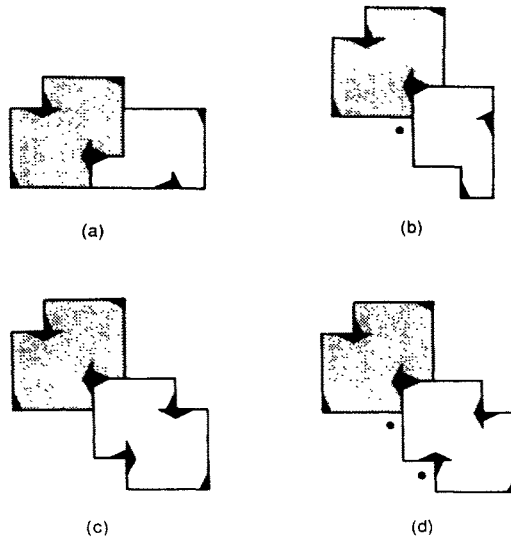
(c)                                   (d)

**Fig. 10**

possibilities for filling both waists, namely with two A-tiles (Figs. 11(a) and (b)), with one A-tile and one B-tile (Figs. 11(c) and (d)), or with two B-tiles (Fig. 11(e)).

If two A-tiles are used, then a B-tile is forced to fit into the position shown by the dashed lines in Figs. 11(a) and (b); these diagrams show the two ways in which such a B-tile can be placed. In the case of Fig. 11(a), either the "nose" of an A-tile or the "rear" of a B-tile can fit into the waist marked ★, yielding either the patch of Fig. 8(a) or, with two extra A-tiles whose positions are forced, the patch of Fig. 8(b). In the case of Fig. 11(b), if an A-tile is fitted into the waist marked by a star, then the tiling cannot be continued (we obtain the arrangement of tiles shown in Fig. 11(d) reflected in a vertical mirror). If a B-tile is fitted with the arrows at its waists pointing downward and to the left, then the tiling cannot be continued. On the other hand, if the arrows at its waists point upward and to the right, then two other A-tiles are forced and we arrive again at the patch of Fig. 8(b). In the case of Fig. 11(c), two A-tiles are forced as shown by the dotted lines, but the tiling cannot be continued since no tile will fit into the reflex angle marked ●. In the case of Fig. 11(d) the two A-tiles that are forced yield the patch of Fig. 8(a). Finally, in the case of Fig. 11(e), four A-tiles are forced and we arrive again at the patch of Fig. 8(b).

To summarize, we see that in any tiling by the set A4, the B-tiles are of two kinds:

   (i) those into whose waists fit two A-tiles; and
   (ii) those into whose waists fits at least one B-tile.

Each B-tile of the second kind can be built up into one of the patches shown in Fig. 8, using either A-tiles, or B-tiles of the first kind. Also, the "nose" of each A-tile must fit into the waist of a B-tile, and thus be part of the patches just
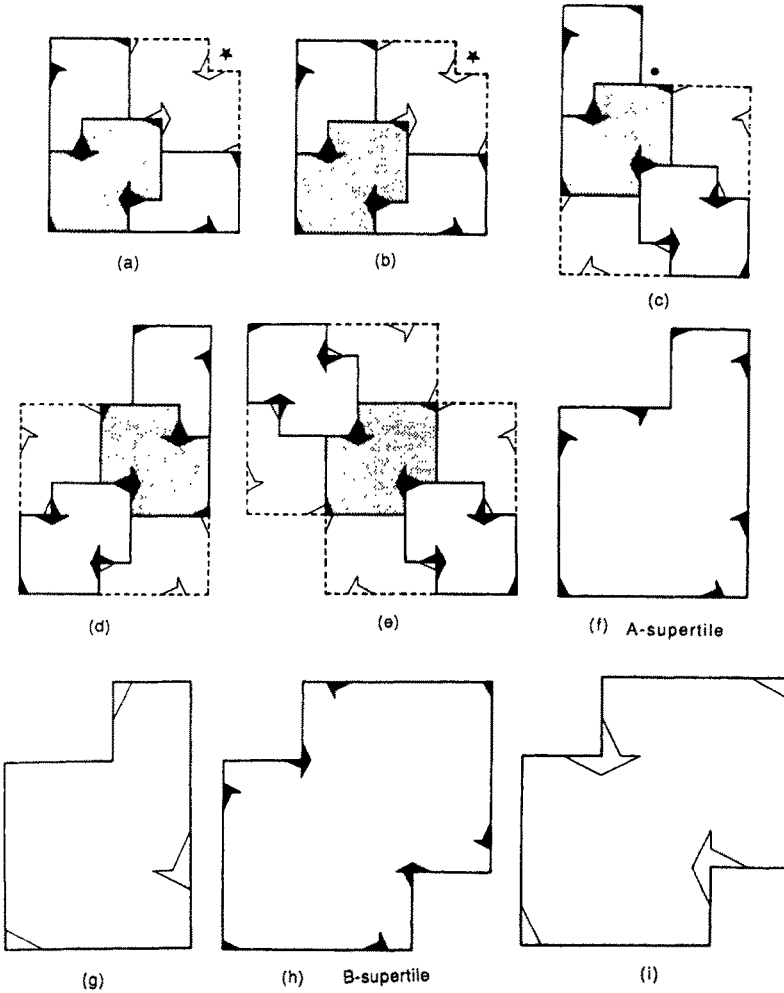
Fig. 11

described. Hence we may weld the tiles in these patches together, to form either an A-supertile (Fig. 11(f)) or a B-supertile (Fig. 11(h)). Moreover, this process is unique, one supertile arising from each B-tile of the second kind. This completes the first part of the proof.

As in the discussion of the set A2, supertiles with markings inherited from the prototiles appear, at first sight, to fit together in more ways than is permitted by enlargements of the original markings, shown in Figs. 11(g) and (i). Some possibilities are indicated in Fig. 12. However, it is easily verified that in each case the tiling cannot be continued. In fact, in the cases shown, no tile will fit into the space marked ●. We deduce that the matching condition for the supertiles is exactly equivalent to that for the original prototiles. This completes the proof of the aperiodicity of the set A4.
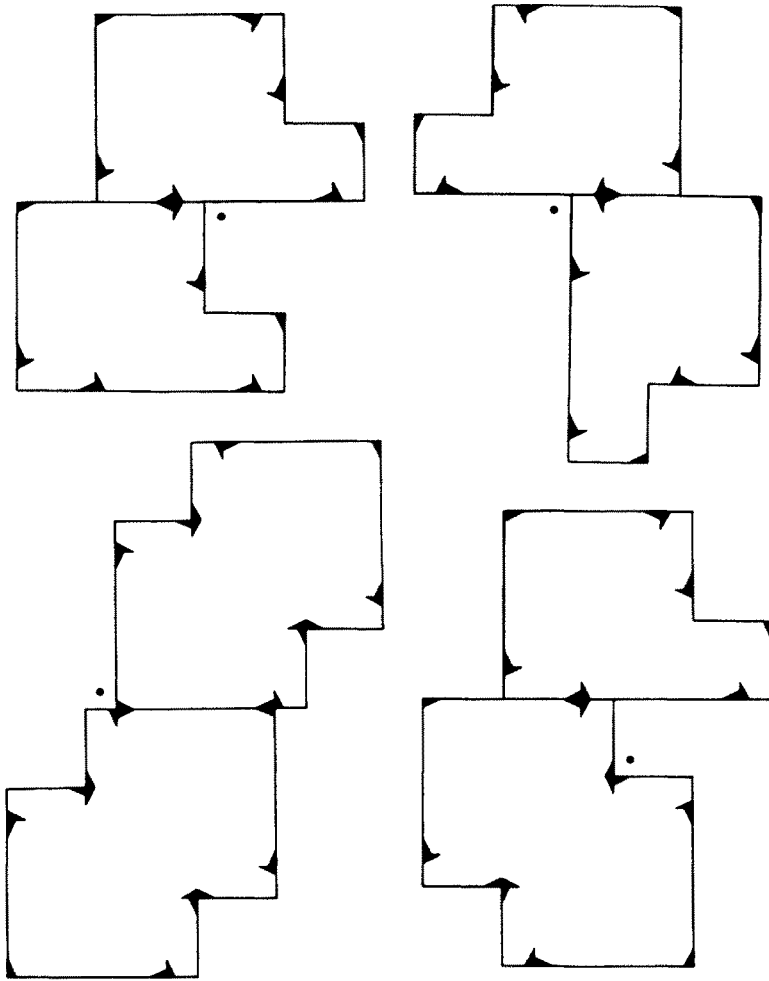
Fig. 12

## 4. The Aperiodic Set A5

The tiles of this set are shown, in three different variants, in Fig. 13. In Fig. 13(a) there are four tiles, two of which are key tiles. In Fig. 13(b) there are only two tiles, one a square and the other a rhomb; instead of key tiles there are markings on these tiles, and there is a matching condition requiring that the quarter circles must pair up to form semicircles, and the angular marks must form "arrows" (or "houses"). The sides of both tiles are equal in length, and the angles of the rhomb are 45° and 135°. The third variant, which is the one most convenient to use in our proof, is shown in Fig. 13(c). In this variant each square tile has been cut into two congruent triangles, and an additional mark (a half arrowhead) has been made on these triangles to ensure that they can fit only by pairing up to form the square
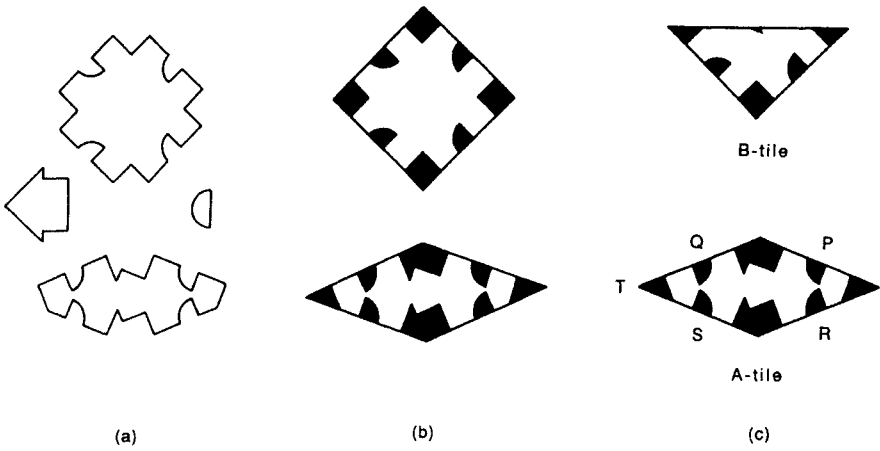
B-tile

Q      P

T

S      R

A-tile

(a)                                 (b)                                 (c)

**Fig. 13**

tiles of Fig. 13(b). We refer to the tiles of the third variant as A-tiles and B-tiles. Also it will be useful from time to time to refer to the square tile in Fig. 13(b) as a 2B-tile.

In Fig. 14 we show an example of a tiling using these tiles, and in Fig. 15 we show how the A-supertile and B-supertile are built up from the original tiles. Our objective is, of course, to establish that in any tiling using these prototiles, composition into the supertiles is unique.

Since it is obvious from examination of the corner markings that there exists no tiling by B-tiles only, we deduce that A-tiles must be present in every tiling.
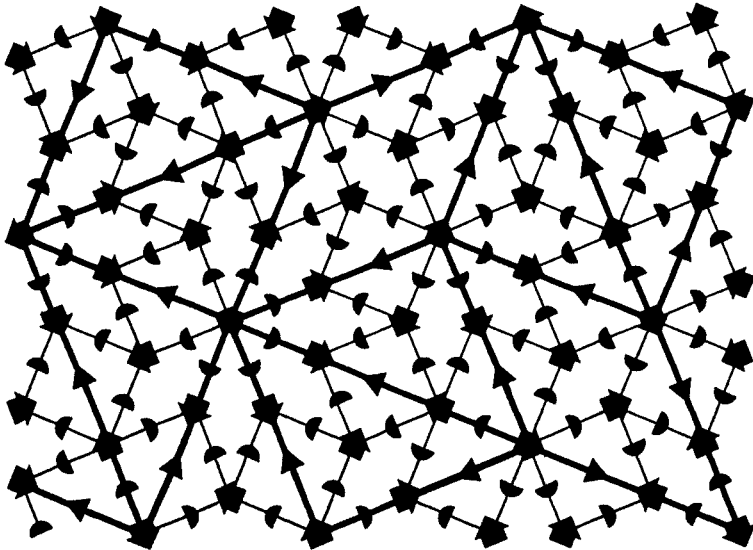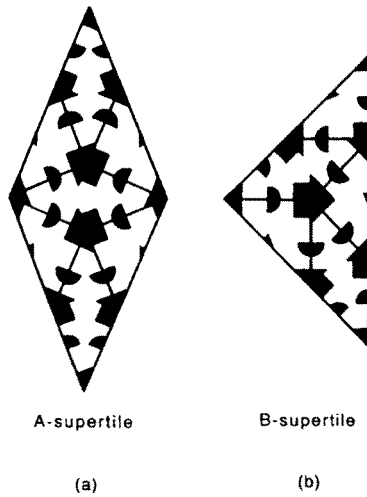


**Fig. 14**

A-supertile          B-supertile

(a)                  (b)

**Fig. 15**

We therefore proceed to analyze possible arrangements of tiles adjacent to an A-tile. If the sides of an A-tile are labeled P, Q, R, and S as in Fig. 13(c), then it is trivial to show that another A-tile can fit along the sides *P*, *R*, but not along sides *Q*, *S* (for in these cases the matching condition at the obtuse corners is violated). On the other hand, a 2B-tile can fit (in two possible ways) along any side *P*, *Q*, *R*, or *S* of an A-tile. It follows that there are just three possible sets of adjacents for an A-tile:

  (i) four B-tiles;
  (ii) three B-tiles and one A-tile; or
  (iii) two B-tiles and two A-tiles.

We examine these possibilities in turn.

  (i) *Four B-tiles.* Consider the right side of the A-tile shown shaded in Fig. 16. The adjacent B-tiles must be oriented as shown—in other orientations no tile will fit into the space marked ●. With the B-tiles in the positions shown, an A-tile will fit into the space ● and moreover can do so in two possible ways shown in Figs. 17(a) and (b). However, the arrangement of Fig. 17(b) cannot be continued to a



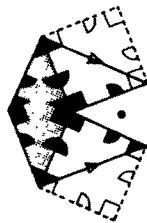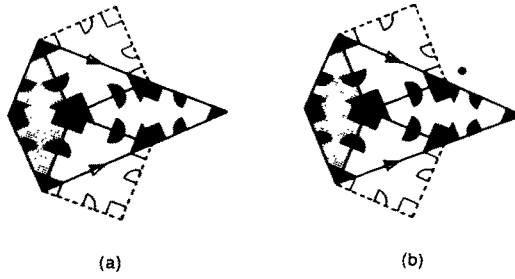**Fig. 16**

(a)                                        (b)

**Fig. 17**

tiling since no tile will fit into the space ●. Hence the tiles must be arranged as in Fig. 17(a). Exactly the same considerations apply to the left side of the A-tile, which must therefore be surrounded by four B-tiles and two A-tiles as in Fig. 15(a). In this way, every A-tile with four adjacent B-tiles leads to a unique A-supertile.

(ii) *Three B-tiles and one A-tile.* One side of the A-tile under consideration is adjacent to two B-tiles, which must therefore be arranged as described in (i). The other side of the A-tile must be adjacent to one A-tile and one B-tile; in view of the preceding remarks these must be arranged as in Fig. 18(a); the only possible variation is that either A-tile may be flipped over about its long diagonal. (To make the diagram applicable in all cases we have omitted the markings at the acute angles of the A-tiles.) The position of the 2B-tile is forced.

We deduce that, depending upon which side of the original A-tile the two B-tiles lie, the tiles must be arranged as in Figs. 18(b) or (c). The only variation possible is that the A-tile on the bottom of Fig. 18(b) may be flipped over about its long diagonal. However, the arrangement in Fig. 18(c) cannot be continued to a tiling
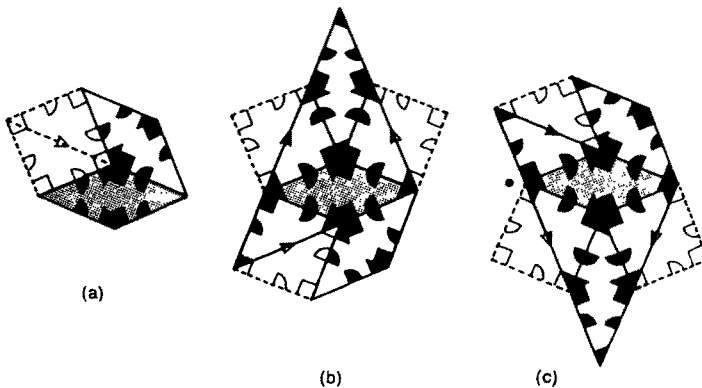


(a)

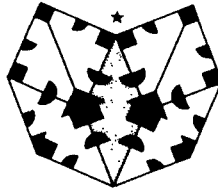(b)                            (c)

**Fig. 18**

**Fig. 19**

since no tiles will fit into the space marked •. We conclude that the tiles must be arranged as in Fig. 18(b). Ignoring the A-tile on the bottom, it will be seen that we have constructed a B-supertile (see Fig. 15(b)). Thus every A-tile with three adjacent B-tiles leads to a unique B-supertile.

(iii) *Two B-tiles and two A-tiles.* Since A-tiles can fit along only two sides of an A-tile (see the earlier remark to this effect), we deduce that in this case the adjacents to an A-tile must be arranged as in Fig. 19. The only possible variation is that each of the A-tiles may be flipped over about its long diagonal, though out of the eight possibilities, two violate the matching condition at the acute angles. (In order to make the diagram appropriate for all cases we have intentionally omitted the markings at the acute angles of the A-tiles.) The positions of the 2B-tiles marked X and Y are forced, and these in turn force another A-tile to fit into the position marked ⋆. This A-tile has two adjacent B-tiles on one side and we know it has at least one on the other side. Hence there are either three or four adjacent B-tiles, and so we can construct an A-supertile as described in (i) or a B-supertile as described in (ii).

We summarize as follows. Say that an A-tile "points toward" its acute corner T (see Fig. 13(c)) because of the parts of the arrows at its obtuse corners. If an A-tile has three or four adjacent B-tiles it leads to a B-supertile or an A-supertile as we have described. If it has exactly two adjacent B-tiles, then it must point toward the obtuse corner of another A-tile and this leads to a supertile which contains the A-tile from which we started. We deduce that *every* A-tile belongs to a uniquely determined supertile. Moreover, since it is impossible for isolated 2B-tiles to occur in a tiling by supertiles, we deduce that *every* tile belongs to a (unique) supertile. This concludes the proof of the uniqueness of composition.

In Figs. 20(a) and (c) we show the two supertiles with the markings inherited from those on the original tiles, and in Figs. 20(b) and (d) the same tiles with markings that are enlargements of those on the original tiles. To complete the proof of aperiodicity it only remains to show that the matching conditions implied by these two alternative sets of markings are exactly equivalent. To do this we observe that one can be changed into the other by the replacement scheme indicated in Fig. 21(a). In the case of edges, the equivalence of the matching conditions is obvious. That the same is true for the acute vertices can be seen from Fig. 21(b).

This completes the proof of the aperiodicity for the set A5.
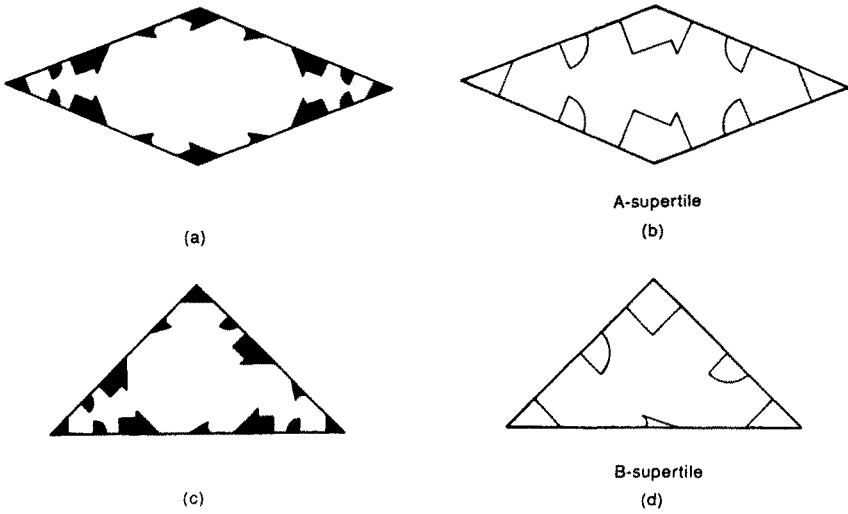
(a)

A-supertile

(b)

(c)

B-supertile

(d)

Fig. 20

## 5. The Aperiodic Set A3

Two variants of this set of three tiles are shown in Fig. 22. In Fig. 22(a) we have three unmarked tiles, and in Fig. 22(b) three tiles for which the matching condition is specified by marks which are parts of circles which must fit together to form complete circles. As is easily seen, the matching conditions in the two variants are not exactly equivalent. In the variant of Fig. 22(b) the tiles can be "turned over,"
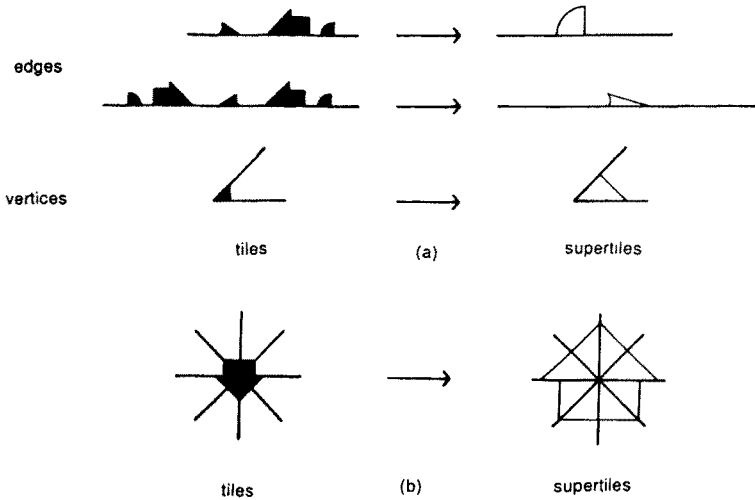


edges

vertices

tiles       (a)       supertiles

tiles       (b)       supertiles

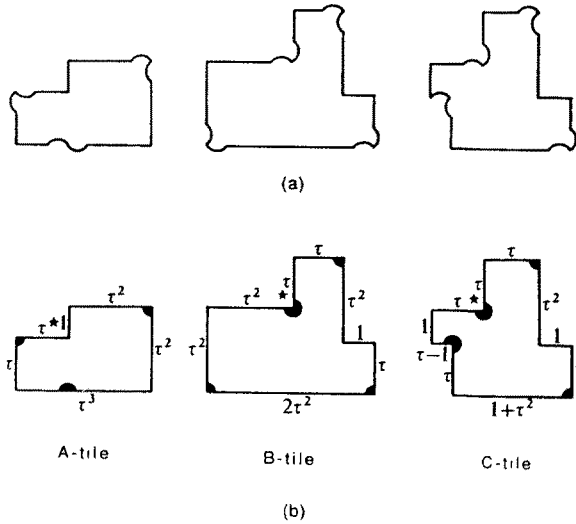Fig. 21

(a)

A-tile        B-tile        C-tile

(b)

**Fig. 22**

that is, their mirror images can be used, whereas the matching conditions in Fig. 22(a) precludes this possibility. Consequently, the tiles of Fig. 22(b) admit a greater variety of tilings than those of Fig. 22(a). In order to simplify the discussion we use the tiles of Fig. 22(b) but impose the additional condition that they *must not be turned over*. This restriction can be achieved with marked tiles if we replace each circle by a marking that differs from its mirror image, for example, by a four-armed swastika. However, we retain the circular markings in the following discussion because occasionally we wish to comment on the additional possibilities that arise if we are allowed to use mirror images of the tiles.

In Fig. 22(b) the dimensions of the tiles are indicated; as before, $\tau = 1.618034\ldots$ is the golden number. Figure 24 shows an example of a tiling by this set of tiles. Here heavier lines divide the tiling into patches of tiles, each of one of the three kinds shown in Fig. 23. When the tiles in each of these patches are welded together we obtain the corresponding supertiles, each of which is $\tau$ times as large as one of the original prototiles.
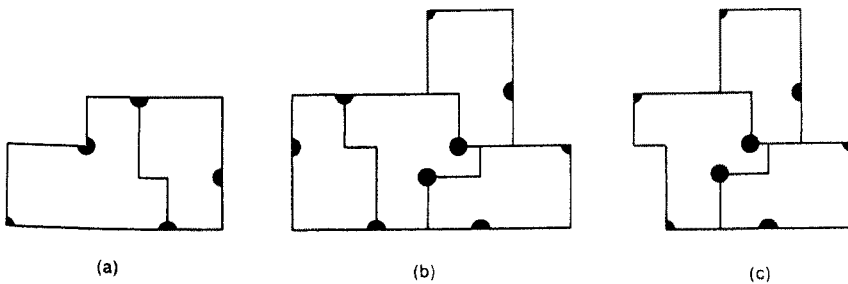


(a)                    (b)                    (c)
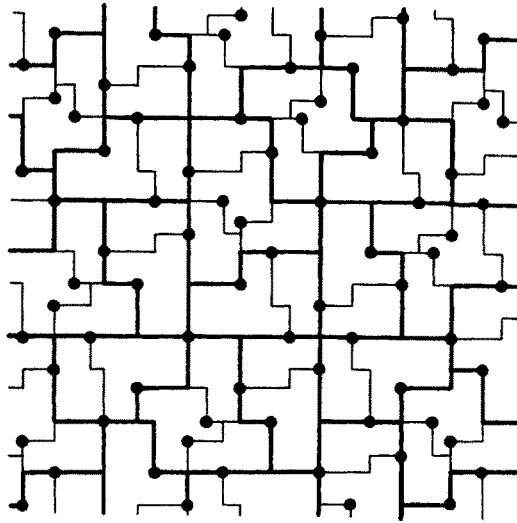
**Fig. 23**

**Fig. 24**

To establish that composition into these supertiles is unique we begin by asserting that no tiling is possible using B-tiles alone. In fact, into the reflex angle of the B-tile marked ★ in Fig. 22(b) must fit a corner marked with a quarter circle. Hence there are just the three possibilities shown in Figs. 25(a)–(c). The arrangement in Fig. 25(a) cannot be continued since no B-tile will fit into the reflex angle marked ●. The reflex angle marked ★ in Fig. 25(b) can be filled in only one way and this leads to the arrangement of four tiles drawn in solid lines in Fig. 25(d). These force the three tiles shown in dashed lines, and the tiling cannot be continued since there is no way to fill the space marked ●. Finally, we are left with the arrangement shown in Fig. 25(c). This implies that all B-tiles are translates of each other, and it is trivial to check that there is no tiling by B-tiles with this property. We deduce that there is no tiling by B-tiles alone. (If mirror images of B-tiles are allowed, the assertion that there is no tiling by B-tiles alone is still correct, but many more possible arrangements of tiles have to be examined.)

Since we now know that every tiling by this set of prototiles must contain an A-tile or a C-tile, we can proceed as follows. Consider which tiles will fit into the reflex angle of a C-tile marked with a star in Fig. 22(b). We must look for a corner of a tile marked with a quarter circle, from which we deduce that there are just seven possibilities:

an A-tile can be fitted in two ways;
a B-tile can be fitted in three ways; and
a C-tile can be fitted in two ways.

(These numbers would be doubled if mirror images of the tiles were allowed.) We do not go into details because it is obvious that in all these cases except one the tiling cannot be continued since there is no tile that will fit into the position near the C-tile marked ● in Fig. 22(b). The only possibility is that shown in Fig. 26(a).

(a)  (b)  (c)

(d)

**Fig. 25**



(a)

(b)

(c)  C-supertile

(d)

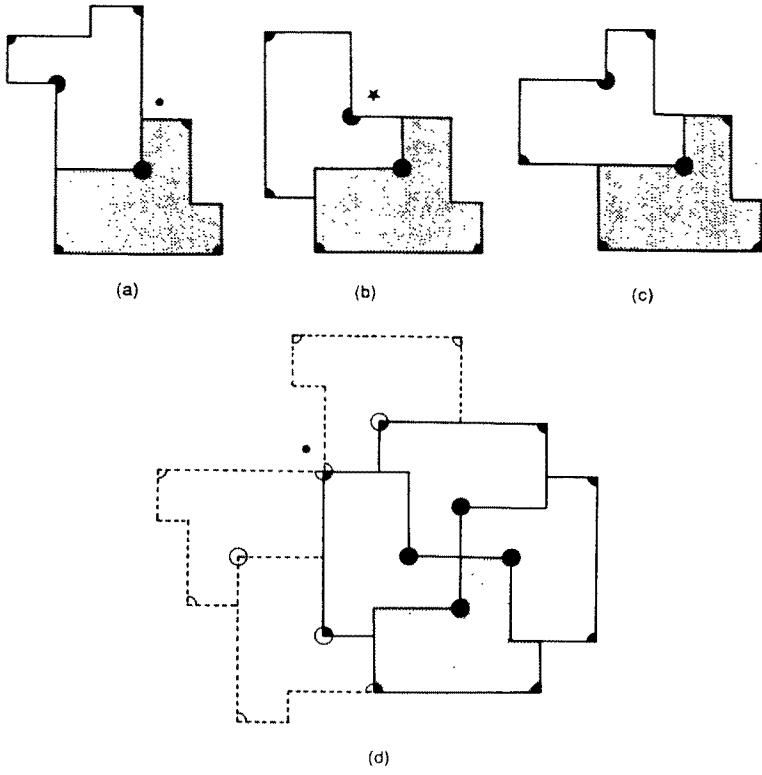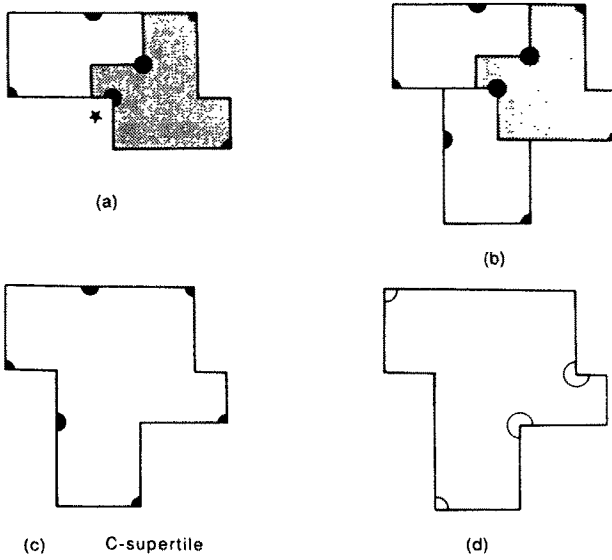**Fig. 26**

Now consider tiles that fit into the reflex angle marked by a star in Fig. 26(a). Here there are six possibilities; each of the tiles A, B, or C can fit in two ways. Again only one of these possibilities leads to a patch of tiles which may be continued, namely that shown in Fig. 26(b). Welding these tiles together we obtain a C-supertile, see Fig. 26(c), and since *every* C-tile is necessarily associated with two A-tiles in this way, we arrive at a tiling by A-tiles, B-tiles, and C-supertiles.

Now consider those tiles which can be fitted into the reflex angle of an A-tile, marked with a star in Fig. 22(b):

an A-tile can be fitted in one way (Fig. 27(a));
a B-tile can be fitted in two ways (Figs. 27(b) and (c)); or
a C-supertile can be fitted in one way (Fig. 27(d)).

However, only two of these arrangements can be continued, namely those shown in Figs. 27(c) and (d). In the other cases it is impossible to fit a tile into the position
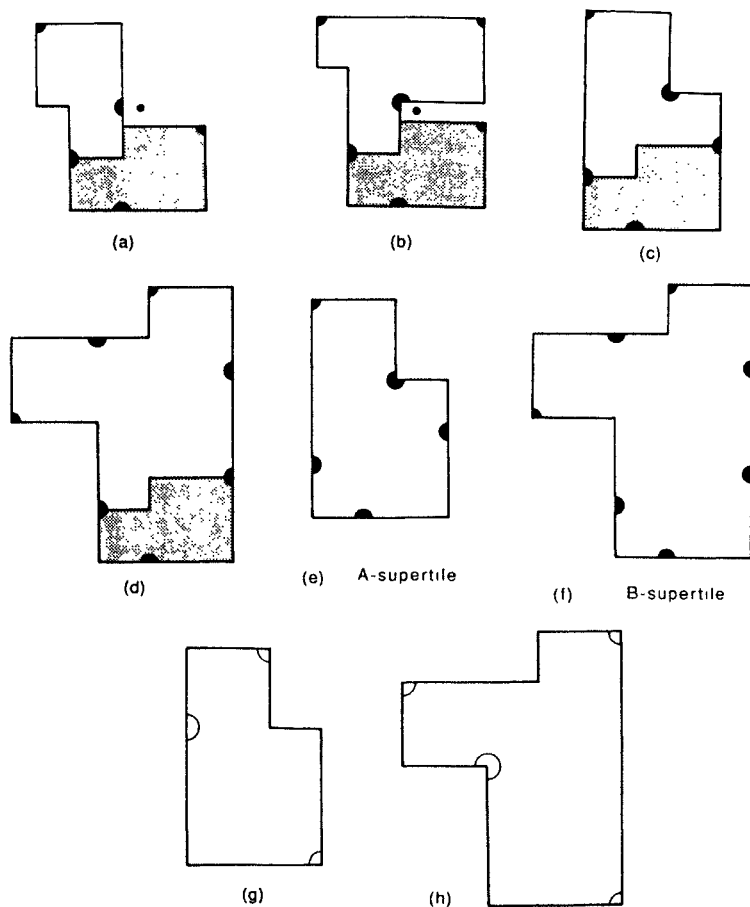


(a)

(b)

(c)

(d)

(e)    A-supertile

(f)    B-supertile

(g)

(h)

Fig. 27

marked ● in the diagrams. Hence we weld together an A-tile to a B-tile as in Fig. 27(c) to form an A-supertile, see Fig. 27(e), and we weld an A-tile to a C-supertile as in Fig. 27(d) to form a B-supertile, see Fig. 27(f). As these are the only possibilities we arrive, in a unique way, at a tiling by A-supertiles, B-supertiles, and C-supertiles. This completes the first part of the proof.

The second part proceeds as in the previous sections. If the A-, B-, and C-supertiles were marked as in Figs. 26(d), 27(g), and 27(h), the matching condition would be precisely the same as with the original prototiles. It therefore only remains to show that the matching conditions specified by the markings in Figs. 26(c), 27(e), and 27(f) are equivalent. Again, at first sight it appears that the latter allow the tiles to be juxtaposed in more ways. However, each additional case can be eliminated by showing that the tiling cannot be continued. This completes the second part of the proof, and establishes the aperiodicity of the set A3 under the restriction that mirror images of the tiles must not be used.

We conclude by considering what may happen if this restriction is removed. The ways in which the tiles can fit together can be analyzed in a similar manner to that described above, except that now many more possibilities need to be considered. Additional tilings can arise in the following manner.

It is possible to construct a tiling of a half-plane $H$, for example in the following way. Start from any tile, enlarge it in the ratio $\tau : 1$ and then cut it up into tiles of the original size, as shown in Fig. 23. This process, known as *decomposition*, may be repeatedly applied to yield arbitrarily large patches of tiles similar in shape to the tile from which we started. Then a slight modification of the *Extension Theorem* [6, Chapter 3] implies that a tiling of a half-plane is possible.

Let the bounding line of the half-plane $H$ be denoted by $L$. Then by reflecting all the tiles of $H$ in $L$ we obtain a tiling of the whole plane in which the tiles and their mirror images occur with equal frequency. We believe that the only tilings by the set A3 in which mirror images of the tiles are used arise in this way, but we cannot prove this conjecture.

## 6. Remarks and Observations

(i) Though the main purpose of the preceding five sections is to prove the aperiodicity of the sets A2, A3, A4, and A5, we have also, incidentally, established the aperiodicity of various other sets of prototiles. For example, in Section 2, in our discussion of the set A2 we showed that the corresponding A-supertile and B-tile also form an aperiodic set. Several similar instances can be found, and others can easily be devised.

(ii) In Chapter 10 of [6] it is shown that every tiling by Penrose kites and darts is *metrically prototile balanced*, which means that it is possible to speak about the average number of kites and the average number of darts per unit area of the plane. Moreover, it is shown that the ratio of these averages is $\tau : 1$ (where $\tau$ is the golden number) for *every* tiling by kites and darts. This fact and the irrationality of $\tau$ lead to an alternative proof of aperiodicity (since in any periodic tiling the ratios of the numbers of tiles of different kinds is necessarily rational).

It seems probable that similar assertions can be made about the Ammann sets discussed in this paper, though we have not carried out all details of the proofs. It is interesting, however, to evaluate the corresponding ratios between the numbers of tiles in each case, on the assumption that the ratios exist and take the same values for *all* tilings admitted by the prototiles of a given set. In the case of the set **A2** the ratio of B-tiles to A-tiles is $\tau:1$ (the same as for the Penrose tiles). For **A3**, the ratio of A-tiles to B-tiles to C-tiles is $\tau^3:\tau:1 = 4.236068\ldots:1.618034\ldots:1$. For **A4** the ratio of A-tiles to B-tiles is $\sqrt{2}:1$, and for **A5** the ratio of A-tiles to 2B-tiles takes the same value. Of course, if the existence of these ratios could be established directly and rigorously then the simple algebraic proof of aperiodicity indicated above could be applied in each case. (It is easy to show that the composition process applicable to each of these tilings implies the irrationality of the ratio of frequencies of various tiles if the existence of a ratio is assumed. However, the existence of the composition process does not seem to imply the existence of a ratio of frequencies in any acceptable sense. Thus the "proofs" of aperiodicity of various tilings based in such irrational ratios, which were given in a number of publications, appear unconvincing. A discussion of this problem is given in [6].)

The fact that the same ratio $\sqrt{2}:1$ occurs for both sets **A4** and **A5** is not surprising in view of the close relationship between tilings by these two sets. (See Fig. 28 where two such tilings are superimposed.)

These observations lead immediately to the question whether there exist sets of aperiodic prototiles associated with irrational numbers other than $\sqrt{2}$ and $\tau$. (It should be noted that not every set of aperiodic prototiles is associated with irrational numbers. As mentioned on pp. 528 and 530 of [6] in the aperiodic set
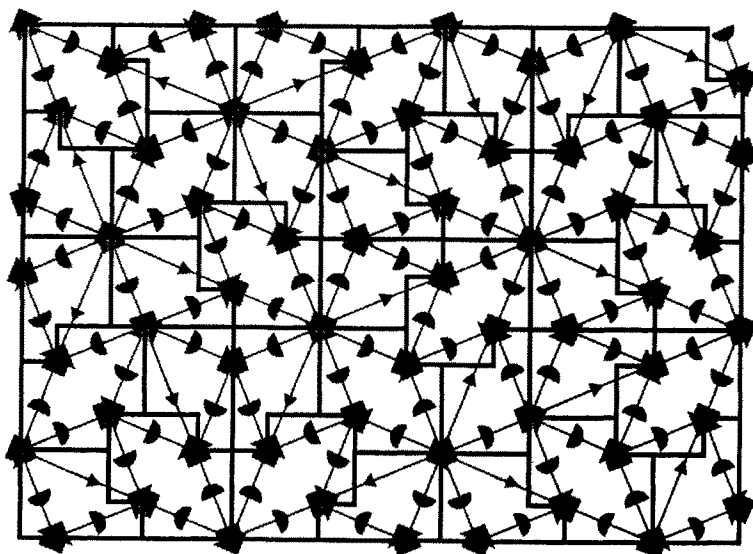


Fig. 28

described by Robinson [14] the ratios of frequencies of the different prototiles are all rational.) We have found aperiodic sets with irrationalities such as $\sqrt{17}$ associated with them, but we have no conjecture concerning the characterization of all numbers that are possible for such ratios.

(iii) Beenker [1] investigated tilings by squares and rhombs which are closely related to the set **A5**. He states [1, p. 55] that for tilings by squares and rhombs it seems impossible to ensure aperiodicity by matching conditions. de Bruijn pointed out to us that this is a confusing statement since what Beenker actually claimed to have proved is that matching rules (and, more generally, any set of rules involving a bounded neighborhood of each rhomb) cannot ensure the property of being the dual of a tetragird. Unfortunately that proof was incorrect, but de Bruijn was able to give a correct one [3]. Beenker's statement may be true as long as we consider only markings or colorings of the edges of the tiles. However, as we have seen in the case of the set **A5**, if we specify a matching condition for the vertices, then it is possible to force aperiodicity. On the other hand, the abstract [7] seems to indicate that weaker matching conditions may also be sufficient. The construction in [18] of nonperiodic tilings by squares and rhombs, with eight-fold rotational symmetry, is of interest since it relies on a composition method, but it has no bearing on the aperiodicity of the set **A5**.

(iv) The proof that suitable sets of aperiodic prototiles can be endowed with matching conditions whose application forces all starting patches to be extendible to (nonperiodic) tilings has been given for Penrose tiles in [11] (see also the popular account in [13]). Similar results for (different) three-dimensional analogues of the Penrose tiles are established in [8] and [4].

(v) Beenker's paper [1] is interesting in that it analyzes tilings by squares and rhombs in detail, and hints at several properties analogous to those known for the Penrose tiles. We mention just one of these, concerning the existence of *cartwheels* (see [5] and Chapter 10 of [6] for a discussion of cartwheels in Penrose tilings). Figure 29 shows two patches of tiles; the union of the (unmarked) tiles in each
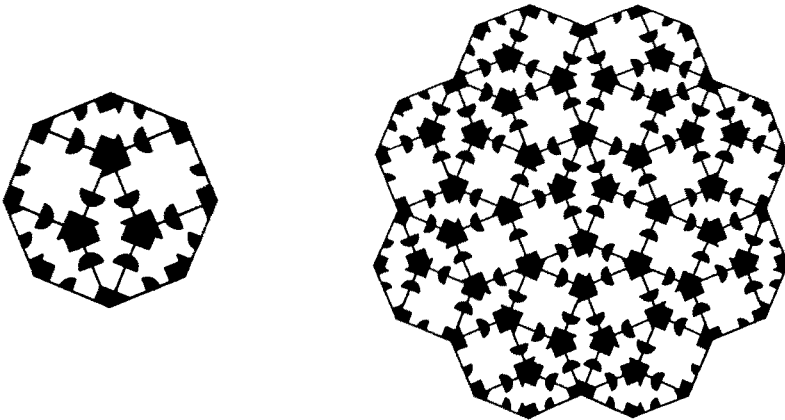


Fig. 29

has symmetry group $d8$. These two patches are called *cartwheels* and are the first two in a sequence of such patches obtained by repeated application of decomposition. Then we assert that *in any tiling by the set* **A5** *of prototiles, every tile lies in a cartwheel of every size.* A proof can be constructed using the processes of composition and decomposition. In addition we remark (compare [1]) that analogues of Conway worms can be defined, and this leads to a theory of holes in tilings by the set **A5**. There are many possible lines of investigation, and we hope that a systematic investigation of these and related properties will be undertaken.

(vi) Following de Bruijn's algebraic approach to Penrose tiles [2], many authors have considered the construction of nonperiodic or "quasi-periodic" tilings as projections or sections of higher-dimensional periodic structures (see, for example, [16] and [17]). These constructions appear to have little—if any—relevance for aperiodic tilings. The reason for this is that such constructions deal only with particular tilings admitted by a given set of prototiles, without considering whether other tilings with the same prototiles are possible. In fact, in most cases it is quite clear that the sets of prototiles are not aperiodic. An insightful discussion of the concepts of aperiodicity and quasi-periodicity, and of the relations between them, is presented in [15], where many references can also be found.

(vii) Many papers dealing with aperiodic sets of prototiles or with quasi-periodic tilings have appeared since the original version of this paper was submitted, and it would lead too far afield to try to present here even a brief account of these developments. In particular, questions regarding the three-dimensional case have received much attention; however, it seems very hard to reach firm conclusions regarding the aperiodic character of most of the sets of tiles investigated (in contrast to the clear possibility of their forming nonperiodic tilings). However, it is exciting to report that an aperiodic set for the three-dimensional space, consisting of a single tile and requiring no adjacency conditions, has been found by Dr. Peter Schmitt.

## References

1. F. P. M. Beenker, Algebraic theory of non-periodic tilings of the plane by two simple building blocks: a square and a rhombus. Technical Report 82-WSK-04, Eindhoven University of Technology, September 1982.
2. N. G. de Bruijn, Algebraic theory of Penrose's non-periodic tilings. *Nederl. Akad. Wetensch. Proc. Ser. A* **84** (1981), 39–66.
3. N. G. de Bruijn, Remarks on Beenker patterns. (Mimeographed notes.)
4. L. Danzer, Three-dimensional analogs of the planar Penrose tilings and quasicrystals. *Discrete Math.* **76** (1989), 1–7.
5. M. Gardner, Extraordinary nonperiodic tiling that enriches the theory of tiles. *Scientific American*, January 1977, pp. 110–121.
6. B. Grünbaum and G. C. Shephard, *Tilings and Patterns.* Freeman, New York, 1986.
7. A. Katz, Construction of matching rules for quasiperiodic tilings. Abstract 88A-51-9, *Abstracts Amer. Math. Soc.* **9** (1988), 437.
8. A. Katz, Theory of matching rules for the 3-dimensional Penrose tilings. *Comm. Math. Phys.* **118** (1988), 263–288.

9. D. Levine and P. J. Steinhardt, Quasicrystals. I. Definition and structure. *Phys. Rev. B* **34** (1986), 596–616.
10. A. L. Mackay, Crystallography and the Penrose pattern. *Phys. A* **114** (1982), 606–613.
11. G. Y. Onoda, P. J. Steinhardt, D. P. DiVincenzo, and J. E. S. Socolar, Growing perfect quasicrystals. *Phys. Rev. Lett.* **60** (1988), 2653–2656.
12. R. Penrose, The role of aesthetics in pure and applied mathematical research. *Bull. Inst. Math. Appl.* **10** (1974), 266–271.
13. I. Peterson, Tiling to infinity. *Sci. News* **134** (1988), 42.
14. R. M. Robinson, Undecidability and nonperiodicity of tilings of the plane. *Invent. Math.* **12** (1971), 177–209.
15. M. Senechal and J. Taylor, Quasicrystals: the view from Les Houches. *Math. Intelligencer* **12** (1990), 54–64.
16. J. E. S. Socolar and P. J. Steinhardt, Quasicrystals. II. Unit-cell configurations. *Phys. Rev. B* **34** (1986), 617–647.
17. P. J. Steinhardt, Quasicrystals. *Amer. Sci.* **74** (1986), 586–597.
18. Y. Watanabe, M. Ito, and T. Soma, Nonperiodic tessellation with eightfold rotational symmetry. *Acta Cryst. Sect. A* **43** (1987), 133–134.