Remarks on Penrose tilings

Remarks on Penrose Tilings

N. G. de Bruijn

Department of Mathematics and Computing Science, Technological University Eindhoven, Eindhoven, The Netherlands

1. Introduction

1.1. This paper will cover some details on Penrose tilings presented in lectures over the years but never published in print before. The main topics are: (i) the characterizability of Penrose tilings by means of a local rule that does not refer to arrows on the edges of the tiles, and (ii) the fact that the Ammann quasigrid of the inflation of a Penrose tiling is topologically equivalent to the pentagrid that generates the original tiling.

The fact that any Penrose tiling is the topological dual of the Ammann quasigrid of the inflation was first noticed by Socolar and Steinhardt ([9]). They presented it in the equivalent form that the topological dual of the Ammann quasigrid of a Penrose tiling is the *deflation* of that tiling.

The Ammann quasigrid of the inflation of a Penrose tiling can also be defined as the union of the central lines of the stacks of that tiling. Therefore I refer to that union as the *central grid* of the tiling. It can be defined independently of the original notion of Ammann grid. Actually the definition of the central grid can also be given for other kinds of tilings where there is no obvious definition of an Ammann grid and no obvious definition of deflation.

The paper is intended to be readable more or less independently of previous ones, at least in the sense that all relevant notions will be explained in the paper itself.

1.2. My geometric terminology will use the notion of shapes and 1-shapes in the plane. A *shape* is an equivalence class of figures under similarity transform. Similarity includes multiplications (with respect to a point), shifts and rotations, but no reflections with respect to a line. In terms of complex numbers, this similarity transform just means linear transformation. A *1-shape* is defined similarly, but without multiplications. That means that similarity is replaced by congruence. In other words, figures with the same 1-shape have the same shape and the same size.

Throughout this paper I shall use the two 1-shapes V and W, pictured in Figure 1.1. They are rhombs, with all edges having unit length. The acute angles of V are 36°, and those of W are 72°. V will be called the *thin rhomb* and W the *thick rhomb*. The word "rhomb" will always refer to a V or a W.

The *arrowed rhombs* V_a and W_a (the subscript a stands for "arrowed") are obtained from V and W by putting single and double arrows on the edges in the way depicted in Figure 1.2. They will be called the *Penrose rhombs*.

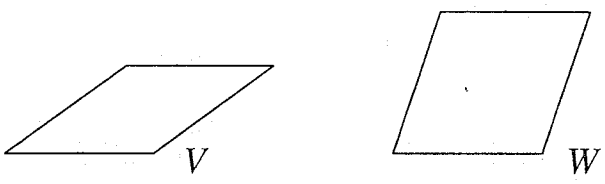

Fig. 1.1. The 1-shapes V (thin rhomb) and W (thick rhomb).

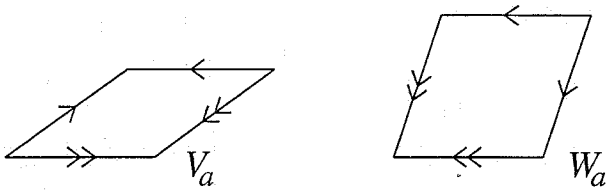

Fig. 1.2. Penrose's arrowed rhombs V_a (thin) and W_a (thick).

1.3. A *Penrose tiling* is a tiling of the plane by V_a's and W_a's with the property that two tiles always have either nothing, or a vertex, or a full edge in common; in the latter case they are called *direct neighbors* and it is required that along the common edge they have the same kind of arrow in the same direction. Figure 1.3 shows a fragment of such a tiling.

Various procedures for obtaining all Penrose tilings are known, like

(i) Penrose's use of deflation, in combination with a (non-constructive) selection argument (see [4] for an exposition).

(ii) the use of deflation, in combination with "updown-generation" (see [5]).

(iii) forming the dual of a "pentagrid" and providing the arrows afterwards (see [1]).

In [5] the relation between (ii) and (iii) was studied in detail.

Somewhat more complicated, but nevertheless interesting and promising, are ways to take the Ammann bars (see [7, 9]) as the basis of the Penrose tilings.

1.4. One of the topics to be treated in the present paper is the question how one can determine whether a tiling of the plane by V's and W's can be provided with arrows so as to form a Penrose tiling with V_a's and W_a's. It

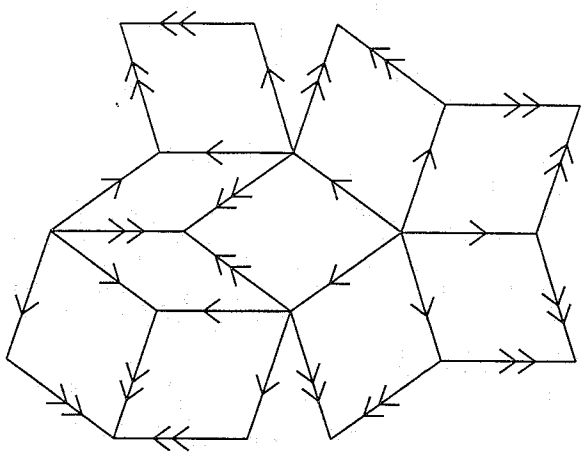

Fig. 1.3. A piece of a Penrose tiling.

will be shown (section 2.6) that this can be settled by inspecting, for every rhomb in the unarrowed tiling, the figure formed by its four direct neighbors.

Section 2.9 will pay attention to the question of a tiling give by the *vertices* only.

2. Arrow-free characterization of Penrose tilings

2.1. The term *unarrowed rhomb tiling* will mean a tiling of the plane with V's and W's, with the condition that neighboring tiles have a full edge in common.

Any Penrose tiling turns into an unarrowed rhomb tiling by omitting the arrows on the edges.

For any unarrowed rhomb tiling, any way to attach an arrow to every edge such that it becomes a Penrose tiling, will be called an *arrowing* of the tiling, and whenever such an arrowing exists the tiling is called *arrowable.* Not every unarrowed rhomb tiling is arrowable. A simple counterexample is the doubly periodic tiling by thick rhombs only.

It will be shown that the condition for an unarrowed rhomb tiling to be arrowable can be put into a form that does not refer to arrows any more.

2.2. Consider a figure formed by an unarrowed rhomb t and four neighboring rhombs t_1, t_2, t_3, t_4, each one of these having one edge in common

with t. It is required that there is no overlap between any two of these rhombs (apart from possible common edges). Let me call such a figure a *cross*.

If to a cross formed by t, t_1, t_2, t_3, t_4 one adds a number of further rhombs $t_5, \ldots, t_k$ which all have a vertex in common with t, such that there is no overlap between any two of $t, t_1, \ldots, t_k$, and such that the areas around the vertices of t are entirely covered, then that figure will be called an *extended cross*.

2.3. An *arrowing* of a cross or an extended cross is a way to attach an arrow (either single or double) to every edge in the figure such that each one of its rhombs becomes a Penrose rhomb.

Not every cross can be arrowed. And a cross that can be arrowed is not necessarily extendable to an extended cross that can be arrowed.

A cross will be called *perfect* if it is extendable to an extended cross that permits an arrowing.

2.4. It is not hard to make a list of all possibilities for perfect crosses. Starting with a thin rhomb t, trying all possible sets of direct neighbors t_1, t_2, t_3, t_4, and investigating whether they have at least one arrowable extension, one finds 6 possibilities, pairwise related by 180° rotation, leading to 3 different 1-shapes P_1, P_2, P_3. Similarly, if starting from a thick rhomb, one gets to 8 possibilities, which are pairwise related by 180° rotation, so there are 4 different 1-shapes P_4, P_5, P_6, P_7.

For the notion "1-shape of a perfect cross" I shall also use the term *neighborhood pattern*. Figure 2.1 gives them all.

P_1 and P_2 form a pair, in the sense that the mirror image of P_1 has the same 1-shape as P_2. In the same sense P_5 and P_6 form a pair. The others, P_3, P_4 and P_7, are symmetric: the axis of symmetry is the short diagonal of t in P_3 and the long diagonal of t in P_4 and P_7 .

2.5. Each one of the seven neighborhood patterns can be arrowed in exactly one way. These arrowed patterns are called $P_{1a}, \ldots, P_{7a}$, and are shown in Figure 2.2.

2.6. At this stage it becomes possible to answer the question which unarrowed rhomb tilings can be arrowed.

I shall say that an unarrowed rhomb tiling satisfies the *cross condition* if for each one of its rhombs the cross formed by that rhomb and its four direct neighbors is one of the seven 1-shapes $P_1, \ldots, P_7$.

Theorem 2.1. *(i) If an unarrowed rhomb tiling of the plane is arrowable then it satisfies the cross condition.*

(ii) If an unarrowed rhomb tiling of the plane satisfies the cross condition then it can be arrowed in exactly one way.

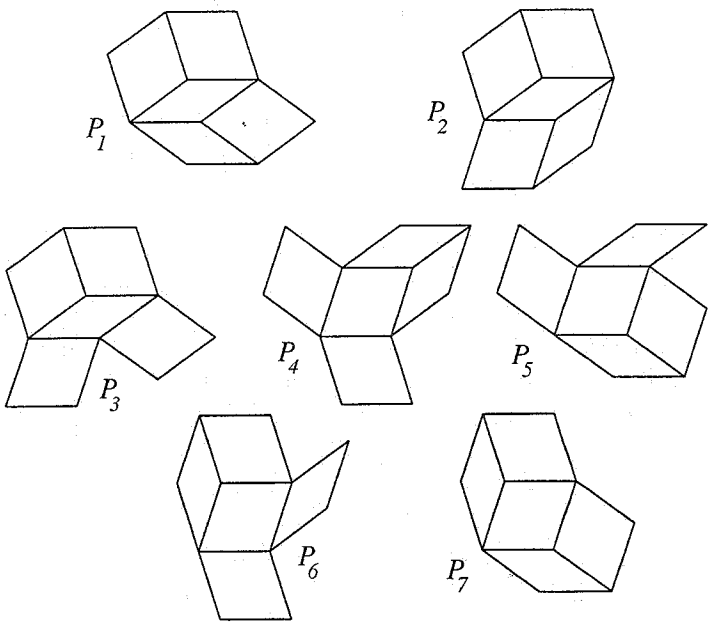

Fig. 2.1. The neighborhood patterns $P_1, \ldots, P_7$.

Proof. (i) This is little more than the fact that the list of neighborhood patterns provided in section 2.4 is exhaustive. If a cross is a part of an arrowable rhomb tiling, then it also has an extension to an extended cross that lies in that tiling, and any arrowing of the whole rhomb tiling implies an arrowing of that extended cross.

(ii) Start from an unarrowed rhomb tiling that satisfies the cross condition. So if t is one of the rhombs in the tiling then it is the central rhomb of a cross of one of the types $P_1, \ldots, P_7$, and that uniquely defines an arrowing of t according to Figure 2.2. So the arrows on t are imposed by the cross of t. They will be called the *imposed arrows* of t.

In this way an arrowing is prescribed for every rhomb in the tiling, but it is the question whether neighbors get the same imposed arrow on their common edge. To that end it has to be shown that if in the tiling two rhombs s, t have an edge in common, then along that edge the arrow of s (imposed by the cross of s) is the same as the arrow of t (imposed by the cross of t). This can be checked without having the full tiling available.

If one of s, t is thin and the other one is thick, then uniqueness of arrowing holds already for the figure consisting of that s and t only. That arrowing on that figure is imposed by the cross of s as well as by the one of t, so in particular the imposed arrow on the common edge is the same in both cases.

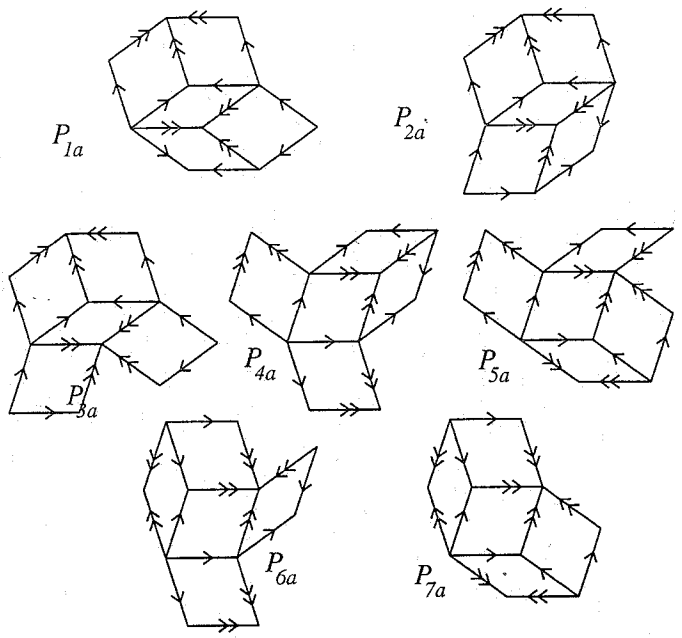

Fig. 2.2. The arrowed neighborhood patterns $P_{1a}, \ldots, P_{7a}$.

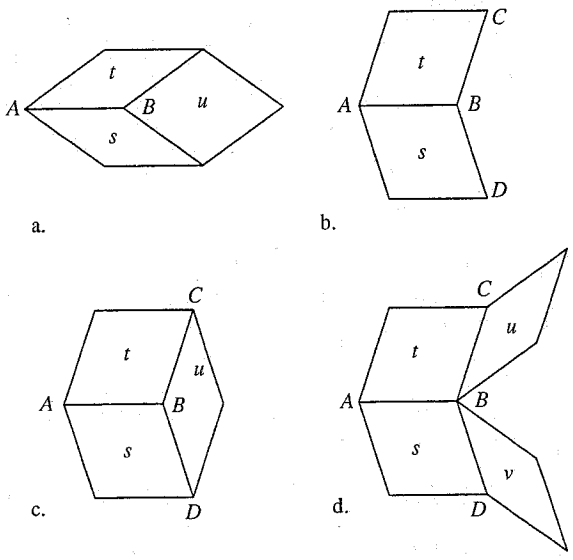

Fig. 2.3. The arrowing of s imposed by the cross of s and the arrowing of t imposed by the cross of t lead to the same arrow on the common edge.

The next case is that s and t are both thin. The cases where the cross of a thin rhomb contains a second thin rhomb are the P_1 and P_2 of Figure 2.1. In both cases s and t have a thick rhomb u as a common neighbor in the way depicted in Figure 2.3a (possibly with s and t interchanged). Just by observing the pair s, u it is seen that the cross of s imposes a double arrow on the common edge of s and t, pointing to the right. With the pair t, u the same result is obtained for the imposed arrow of t.

The remaining case is that both s and t are thick. This is shown in Figure 2.3b (possibly with s and t interchanged). It is a part of a tiling in which the cross condition holds for both s and t. Since the cross of s contains t it cannot contain a thick rhomb u pasted to the edge BC: such a configuration formed by s, t, u does not occur in any of the crosses of s shown in Figure 2.1. The same argument shows that the tiling does not contain a thick rhomb pasted to BD.

The only possibilities to paste thin rhombs to BC and BD are shown in Figures 2.3c and 2.3d.

In Figure 2.3c, the only possible arrowing of the pair s, u has a double arrow from B to A, whence that double arrow is imposed by the cross of s. The same argument applied to the pair t, u leads to the same imposed arrow of t.

In Figure 2.3d the pair t, u shows that the imposed arrow of t on AB is a single one, pointing to the right. Inspection of the pair s, v shows the same imposed arrow of s on AB.

This completes the proof of the theorem.

2.7. As a corollary of the theorem of section 2.6 it can be noted that an unarrowed rhomb tiling can be arrowed in at most one way. This does not require the whole theorem: uniqueness of arrowing holds already for all arrowable extended crosses.

2.8. Every arrowable tiling contains infinitely many copies of each one of $P_1, \ldots, P_7$. Inspection of an arbitrary Penrose tiling shows occurrences of all these perfect crosses, and it is known that any finite configuration occurring in any Penrose tiling occurs infinitely often in any other one.

2.9. An unarrowed rhomb tiling is completely determined by its set of vertices, irrespective of whether the tiling is arrowable or not. In order to see this, it suffices to check that in an unarrowed rhomb tiling any two vertices have distance 1 if and only if they are connected by an edge of a rhomb of the tiling.

It is also easy to see that two vertices in an unarrowed rhomb tiling have distance beteen 0 and 1 if and only if they are the end-points of the short diagonal of a thin rhomb.

The shortest distance exceeding 1 is 2 sin 36° (=1.17557). It is the length of the short diagonal in a thick rhomb, but this distance may occur in a different way between vertices in a pair of adjacent thin rhombs.

3. Deflation and inflation

3.1. This section gives some information about inflation and deflation of Penrose tilings. It is intended as a kind of motivating background, but will not be used in the rest of the paper.

Penrose's idea of inflation and deflation was very central in the discovery of his arrowed rhomb tilings. *Deflation* is a certain way to get from a Penrose tiling to a new tiling with smaller pieces ($\frac{1}{2}(-1+\sqrt{5})$ times the original ones), and *inflation* is the inverse operation, leading to a tiling with bigger pieces. For a description of the process see [1, 6, 7, 8]. It is somewhat easier to describe the deflation as an operation on rhombus *halves*. The half thin rhomb is obtained by cutting a thin rhomb along its short diagonal, the half thick rhomb by cutting a thick rhomb along its long diagonal. For these rhomb halves the deflation is a matter of *subdivision* (see [5]). This is shown in Figure 3.1.

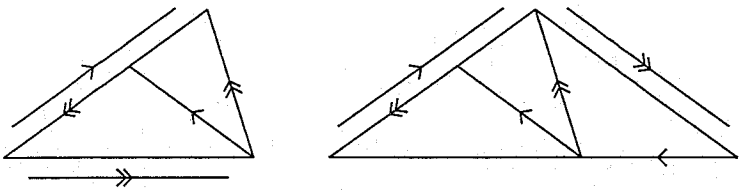

Fig. 3.1. The deflation of the half thin and the half thick rhomb. The full figure on the left is a half thin rhomb, with arrows drawn at the outside. It is subdivided into a half thin and a half thick rhomb, with arrows indicated along the edges themselves. These two small pieces belong to the deflation. Similarly the full figure on the right is a half thick rhomb, and here the deflation has three pieces: a half thin rhomb and two half thick rhombs. There are of course two shapes of half thin rhombs, which are each other's mirror image, and similarly there are two shapes of half thick rhombs. For the mirror images of the half rhombs shown in the figure one has to take the mirror image subdivision.

When this subdivision operation is carried out for every half thin rhomb and every half thick rhomb of a Penrose tiling, the smaller pieces fit together to a Penrose tiling with smaller pieces.

In section 2.3 of [5] it is explained that the idea of inflation comes naturally by observing how in an arrowed rhombus tiling pieces can be grouped

together. Those groups form the pieces of the inflation.

3.2. If deflation is applied to a tiling of a *finite* portion of the plane, and if the result is enlarged by a factor $\frac{1}{2}(1+\sqrt{5})$ afterwards, it becomes a tiling of a bigger part of the plane with pieces of the original size. Infinite repetition of this process and application of an infinite selection principle leads, albeit in an inconstructive way, to Penrose tilings of the full plane (see [4, 6, 8]). But the idea of inflation and deflation can also be used in a completely different and very constructive way for the production of *all* infinite Penrose tilings. The method is *updown generation*, a process that is controlled by taking an arbitrary infinite path in a particular finite automaton (see [5]).

4. Duality

4.1. The key method in [1] was the production of tilings by means of their topological dual. I spend a few words here as a short independent introduction.

4.2. In any Penrose tiling the rhombs are arranged in what I shall call *stacks*. Consider an arbitrary edge e of an arbitrary rhomb r of the tiling. Two rhombs of the tiling are called *e-neighbors* if they have a common edge that is parallel to e. Two rhombs of the tiling are called *e-related* if they can be connected by a chain of e-neighbors. The stack generated by r and e is the set of all rhombs which are e-related to r.

Every rhomb of the tiling belongs to exactly two stacks, and any two stacks have exactly one rhomb in common, unless their e's are parallel.

The idea of a stack can at once be generalized to tilings by parallelotopes in higher dimensional spaces. In that case there can be stacks of various dimensions. In the two-dimensional case the stacks are one-dimensional, but in three dimensions there are two-dimensional stacks that can be considered as unions of one-dimensional stacks.

4.3. Consider a Penrose tiling and a stack generated by some r and e. In any rhomb of the stack connect the mid-points of the edges parallel to e by a straight line segment. These line segments form an infinite broken line that can be called the *back bone* of the stack.

The union of the back bones of all stacks is called the *skeleton* of the tiling. In a skeleton one has *curves* (the back bones), *points* (the intersection points of the curves) and *mazes* (the connected components of the complement of the union of all curves). These curves, points and mazes can be deformed topologically without disturbing their relation to the tiling. On any one of the curves the order of the intersection points with other curves keeps reflecting the order in which the corresponding stack is intersected by other stacks. The topological deformations do not disturb the duality between skeleton and

tiling. In that duality curves correspond to stacks, points to rhombs, mazes to rhomb vertices.

4.4. If some topological deformation of a skeleton is given, one is not yet able to reconstruct the tiling. First one has to know, for any curve of the skeleton, the direction and the length of the edge e that was used in the discussion of the corresponding stack. Once all these are known, the Penrose tiling is determined up to a parallel shift in the plane. That is, the Penrose tiling deprived of its arrows. But the latter is no great loss: since one started from a Penrose tiling in the first place, one knows that a consistent system of arrows exists, and section 2.7 takes care of the uniqueness of the arrowing.

4.5. In [1] it was shown that the skeletons are topologically equivalent to so-called pentagrids, but this had to be taken with a grain of salt in so-called singular cases. The pentagrids are parametrized by real numbers $\gamma_0, \ldots, \gamma_4$ satisfying $\gamma_0 + \ldots + \gamma_4 = 0$. If j is one of the numbers $0, \ldots, 4$ and $\zeta = e^{2\pi i/5}$ then the set of all complex numbers z satisfying $\mathrm{Re}(z\zeta^{-j}) + \gamma_j \in \mathcal{Z}$ is a grid of parallel lines in the complex plane. The superposition of these five grids is called the *pentagrid* generated by $\gamma_0, \ldots, \gamma_4$. The pentagrid is called *singular* if it contains three lines passing through one point.

Key results of [1] are the following. Any non-singular pentagrid is the dual of a Penrose tiling. The singular pentagrids do not have a dual in the ordinary sense, but by infinitesimal perturbations they turn into grids that do have a dual. There can be various perturbations with different effect, and the result is that a singular grid corresponds either to 2 or to 10 different Penrose tilings. With these extra arrangements for the singular cases, *all* Penrose tilings are obtained from pentagrids.

The perturbations will get some more attention in section 6.2 below.

4.6. The idea of producing a tiling as a topological dual of the superposition of a number of parallel grids can be extended to general classes of tilings of spaces of arbitrary dimension by means of parallelotopes. See [2] for a proof that under fairly general conditions the dual covers the whole space uniquely.

4.7. It can be concluded from section 4.5 that the skeleton of a Penrose tiling is topologically equivalent to some regular pentagrid or to some perturbed singular pentagrid. But remarkably, there is always a second and completely different straight line grid with the topological structure of the skeleton. This matter will be treated in sections 5 and 6, where it is also shown that there are not just these two: there are infinitely many straight line grids with the same topology. The pentagrid can be deformed continuously in such a way that it always keeps the same topology and always remains a superposition of five straight line grids, with the central grid as final result. And one can even get beyond that.

5. Ammann bars and central bars.

5.1. An attractive way to introduce the *Ammann bars* of a Penrose tiling is to consider both the thin and the thick arrowed rhomb as a billiard table and to choose particular billiard ball tracks on them. These tracks can already be shown on the rhomb halves, displayed in Figure 5.1. The tracks have been chosen in such a way that when fitting pieces together the segments combine to infinite straight lines. These are called the Ammann bars after their discoverer R. Ammann (see [7, 9]).

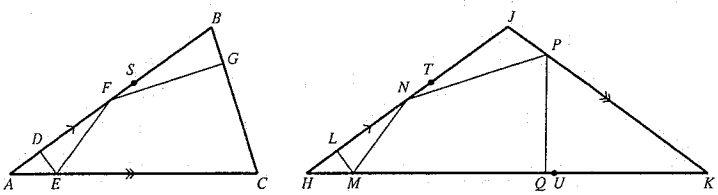

Fig. 5.1. The billiard ball tracks on the arrowed half rhombs. The angles at D, G, L and Q are 90°. The lines EF and FG form equal angles with AB, which is expressed by saying that the billiard ball bounces at the edge AB. At E, M, N and P there are similar cases of bouncing. Finally $AD = HL$, $AF = HN$ and $JP = AE = HM$. In the mirror images of these half rhombs one of course takes the mirror image tracks.

The grid formed by these bars, called *Ammann quasigrid* in [9], has some of the properties of the *pentagrid* used to produce the Penrose tiling in [1]. Both are superpositions of five parallel grids, making angles which are multiples of 72°. In the pentagrid the parallel grids are all equidistant, but in the Ammann quasigrid they show two different distances. On the other hand, the Ammann quasigrid looks much simpler than the pentagrid since its meshes show only a small finite number of shapes. A pentagrid contains infinitely many different mesh shapes, and arbitrarily small ones.

It was observed by J.E.S. Socolar and P.J. Steinhardt (see [9]) that dualization of the Ammann quasigrid of a Penrose tiling leads exactly to the deflation of that tiling.

A quite simple suggestive argument can be given for this statement, although it might not be easy to turn it into a formal proof. It is explained by Figure 5.2. The straight line billiard track segments can be continuously deformed into the segments in Figure 5.3. The points D,E,G,L,M,P can stay

where they were, but F, N and Q move. In the mirror images of these half rhombs the corresponding things are done correspondingly. If in each rhomb of a Penrose tiling the original billiard ball track is deformed continuously into the new track, then the union of all segments shows the picture of a grid deforming itself by bending its bars, all the time keeping the same topology. In the final situation (the one of Figure 5.2) each mesh contains exactly one vertex of the deflation, and that is exactly the one produced by the dualization.

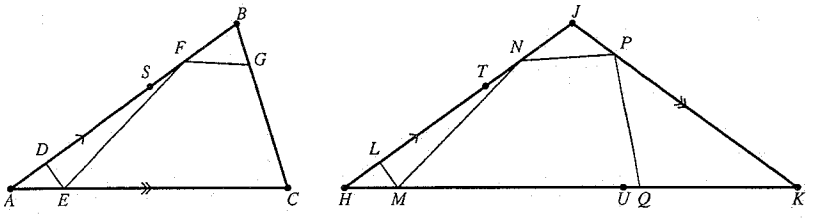

Fig. 5.2. The deformed billiard ball tracks on the half thin and the half thick rhomb. In the mirror images of these half rhombs mirror image tracks have to be taken. The vertices of the deflation in the above half rhombs are A,B,C,S,H,K,J,U,T. Having these tracks in all half rhombs of a Penrose tiling results in a set of mazes where each maze contains exactly one deflation vertex, and in the duality that vertex is exactly the vertex corresponding to that maze.

5.2. The Socolar-Steinhardt phenomenon can also be expressed like this: the Ammann bars of the *inflation* of a Penrose tiling form a quasigrid that is the topological dual of that tiling itself. Let me call those Ammann bars of the inflation the *central* bars of the tiling itself. It will be shown that they can be defined independently of the billiard ball construction, and that a proof for the duality can be given without reference to the argument of section 5.1.

The name "central" was chosen because of the the central position of these bars in the stacks of the tiling (see section 6.4).

The central bars can be found by comparing Figures 3.1 and 5.1. The double arrows of the deflation are from S to A, from C to B, from T to H and from U to J. These are all intersected orthogonally by the billiard ball track. The distance from the intersection point to the end-point of the double arrow is just one fourth of the length of the arrow. This shows that the central bars of a Penrose tiling have to contain the segments EF, GH, TU, VW indicated in Figure 5.3.

It has to be admitted that some of the rhombs contain further pieces of central bars, but in order to build the full central grid it suffices to consider those pieces of Figure 5.3. Since

$$EB = GB = TP = VP, \quad CF = AH = RU = RW \tag{5.2.1}$$

it is obvious that in a stack the segments in the direction of the stack fit together to a full straight line. It is easy to show directly that (5.2.1) implies $EB = GB = TP = VP = \frac{1}{4}$, $CF = AH = RU = RW = \frac{1}{4}(2+\sqrt{5})$, and therefore the central bars can be introduced without any reference to the billiard ball tracks.

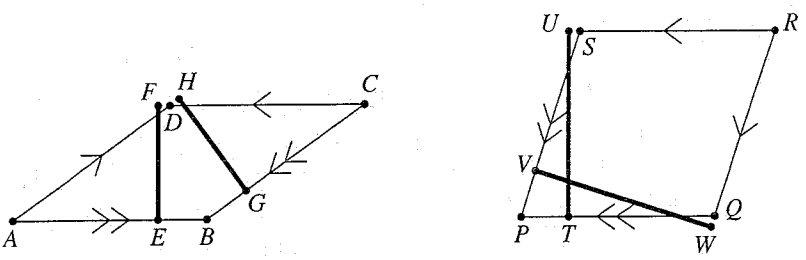

Fig. 5.3. The basic segments of the central bars. In the thin rhomb $ABCD$ (with double arrows AB and CB) the segments are EF and GH, in the thick rhomb $PQRS$ (with double arrows QP and SP) they are TU and VW. There are right angles at E, G, T and V. The points F, H, U, W lie on the extensions of the segments CD, AD, RS, RQ, respectively. The lengths of the bar segments are $EF = GH = \sin 36°$, $TU = VW = \sin 72°$. The positions are determined by $EB = GB = TP = VP = 0.25$. Moreover $DF = DH = SU = QW = (-2+\sqrt{5})/4 = 0.059017$.

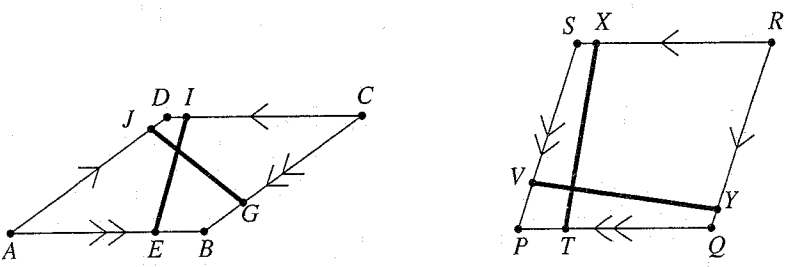

Fig. 5.4. The segments of the deformed bars EI, GJ in the thin rhomb and TX, VY in the thick rhomb, determined by $DI = DJ = SX = QY = 0.1$. Any maze formed by these deformed bars contains exactly one vertex of the Penrose tiling, and that is exactly the vertex corresponding to the maze in the duality.

The segments of Figure 5.3 can be transformed into those of Figure 5.4 by continuous deformation. Taking the segments of Figure 5.4 in all rhombs of a Penrose tiling one gets something that is topologically equivalent to the

skeleton of the tiling (that was defined in section 4.3 by connecting mid-points of the edges).

The transition from Figure 5.3 to Figure 5.4 does not affect the topology of the grid. This is not completely trivial, since (in contrast to what what was described in section 5.1) the operations take place partly *outside* the rhombs. So it has to be made sure that the corresponding operations in neighboring rhombs do not interfere.

In section 6 the topological equivalence of pentagrid and central grid will be shown in quite a different way.

6. Algebraic proof of the topological equivalence

6.1. The various grids to be considered are generalizations of the one described in section 4.5. They are characterized by real numbers $\omega_{j,k}$, where j runs through the set $\{0,1,2,3,4\}$ and k through the set $\mathcal{Z}$ of all integers. The grid determined by these ω's is the superposition of $\Gamma_{\omega,0},\ldots,\Gamma_{\omega,4}$, where each $\Gamma_{\omega,j}$ is is a grid of parallel lines in the complex plane. The line with index k in $\Gamma_{\omega,j}$ is the set of all complex numbers z for which $\mathrm{Re}(z\zeta^{-j}) = \omega_{j,k}$ (as before, $\zeta = e^{2\pi i/5}$).

So the pentagrid described in section 4.5 with its five real parameters $\gamma_0,\ldots,\gamma_4$ satisfying $\gamma_0+\ldots+\gamma_4 = 0$ is the case where $\omega_{j,k} = k - \gamma_j$.

6.2. In [1] it was indicated how the pentagrid defines a Penrose tiling. Singular grids cannot be used directly: they first have to get an infinitesimal deformation in order to admit proper dualization (see [1], section 12).

A few words may be devoted to the meaning of the word "infinitesimal" here. If in a singular pentagrid a grid line, a vertical one, say, is moved over a small finite distance to the right, then triple intersection points may have been avoided in a big finite range only. For a given line in a pentagrid and a given large positive number R there exists a small positive ε such that within the circular disk given by $|z| < R$ all triple intersections are avoided by shifting the line over a distance δ with $0 < \delta < \varepsilon$. But outside the range R such shifts may cause other intersection points, and alter the topology. The topology of the grid obtained by an infinitesimal shift to the right can be defined as the limit of the topology obtained within $|z| < R$ by means of sufficiently small shifts of the grid line.

But of course, perturbation of a pentagrid is not just a perturbation of a single line, but of the set of the five γ's. I now introduce an index p that describes which infinitesimal perturbation has to be taken. It has ten possible values: $0,\ldots,9$, but in the majority of the singular cases there are only two that have a different effect, and in the regular cases all ten have the same effect. In [1], section 9, it was explained that the complex number $\xi = \sum_{j=0}^{4}\gamma_j\zeta^{2j}$ is an essential parameter for the Penrose tiling. It is related to the real numbers μ_j, to be used in this section, given by

$$\mu_j = \gamma_j - (\gamma_{j-1} + \gamma_{j+1})\tau^{-1},$$

for $j = 0, \ldots, 4$, where, as before, $\tau = 2\mathrm{Re}\ \zeta = \frac{1}{2}(-1+\sqrt{5})$, and j is taken mod 5 (so $\gamma_5 = \gamma_0$, etc.). The μ's can be derived from ξ by $\mathrm{Re}(\xi\zeta^{-2j}) = (1-\frac{1}{2}\tau)\mu_j$.

It was shown in [1], section 11 that a pentagrid is singular if and only if for one of the j's there is an α of the form $(1-\zeta)(n_0+n_1\zeta+n_2\zeta^2+n_3\zeta^3+n_4\zeta^4)$ with integers $n_0, \ldots, n_4$ such that $\mathrm{Re}((\xi-\alpha)\zeta^{-j}) = 0$. It is not hard to show that this condition is equivalent to the following one: for at least one value of j there is an integer k such that $(k-\mu_j)\tau$ is an integer.

Now consider infinitesimal perturbations of ξ, given by a positive infinitesimal $\mathrm{d}w$ and a perturbation parameter p (one of the values $0, \ldots, 9$). The perturbation of ξ is $\mathrm{d}\xi = e^{p\pi i/5}\mathrm{d}w$. And perturbations of the γ's that produce this $\mathrm{d}\xi$ can be taken as (cf. [1], formula (9.2)): $\mathrm{d}\gamma_j = (2/5)\mathrm{Re}(\zeta^{-2j}\mathrm{d}\xi)$. The perturbations of the μ's turn out to be

$$\mathrm{d}\mu_j = (1-\frac{1}{2}\tau)^{-1}\mathrm{Re}(e^{(p-4j)\pi i/5})\mathrm{d}w.$$

6.3. The lines of a pentagrid correspond to stacks of the Penrose tiling, and for each stack there is a central bar according to section 5. These central bars can be evaluated in terms of the μ's. This will be explained in section 6.4. The result is that the line with index k in the j-th subgrid of the pentagrid leads to a central bar given by $\mathrm{Re}(z\zeta^{-j}) = \beta_{j,k}$, where, at least in the case of a non-singular grid, $\beta_{j,k} = (1+\frac{1}{2}\tau^{-1})(k+\tau\lceil(k-\mu_j)\tau\rceil-\frac{1}{2}\tau)$. ($\lceil x\rceil$ is the usual notation for the least integer $\geq x$). In singular cases the value of $\lceil(k-\mu_j)\tau\rceil$ can be affected by perturbation of μ_j if $(k-\mu_j)\tau$ is an integer. The effect can be described by means of the notations $\lceil x\rceil_+$ and $\lceil x\rceil_-$, intended as $\lceil x+\mathrm{d}w\rceil$ and $\lceil x-\mathrm{d}w\rceil$, respectively, where $\mathrm{d}w$ is a positive infinitesimal. This just means

$$\lceil x\rceil_+ = \lfloor x\rfloor + 1, \quad \lceil x\rceil_- = \lceil x\rceil$$

for all real values of x. With this notation the result for the central bar becomes

$$\beta_{j,k} = (1+\frac{1}{2}\tau^{-1})(k+\tau\lceil(k-\mu_j)\tau\rceil_{\varphi(p,j)} - \frac{1}{2}\tau), \tag{6.3.1}$$

where $\varphi(p,j)$ stands for $+$ or $-$ according to whether $\mathrm{Re}(e^{(p-4j)\pi i/5})$ is < 0 or > 0.

In the non-singular cases the $\varphi(p,j)$ can be ignored, of course. But it should be noted that it is not so easy to see from the γ's whether a case is singular or not. If one does not want to bother one might just take the $\varphi(0,j)$ as a standard, but it is dangerous to omit the $\varphi(p,j)$ altogether. In the case $\gamma_0 = \cdots = \gamma_4 = 0$ with its ten different perturbations the formula $\beta_{j,k} = (1+\frac{1}{2}\tau^{-1})(k+\tau\lceil k\tau\rceil - \frac{1}{2}\tau)$ would definitely *not* represent the central bars of a Penrose tiling.

6.4. Here are some details about the derivation of the expression (6.3.1) for the central bar of the stack corresponding to a given line of a pentagrid. For simplicity, non-singularity will be assumed. Moreover, the (unessential) restriction is made that $j = 0$, which makes it possible to talk in terms of left and right. Moreover, the letter j becomes available for other purposes.

So the grid line is (with some integer k) $\mathrm{Re}(z) = k - \gamma_0$. The meshes directly to the left and directly to the right of this line produce the vertices of the rhombs of the stack. According to [1], section 5, the vertices are derived from the meshes as follows. Take any point z in the mesh, and form the integers $K_j(z) = \lceil \mathrm{Re}(z\zeta^{-j}) + \gamma_j \rceil$; then the vertex is $\sum_{j=0}^{4} K_j(z)\zeta^j$. With a real variable t the points of the grid line are represented as $(k - \gamma_0) + it$. So the vertices corresponding to the meshes directly to the left of the line are given by $M(t) = k + \sum_{j=1}^{4} \lceil u_j \rceil \zeta^j$, where $u_j = \mathrm{Re}(((k - \gamma_0) + it)\zeta^{-j}) + \gamma_j$. It will be shown that the real part of $M(t)$ takes only four different values if t varies. That real part is $k + (\lceil u_1 \rceil + \lceil u_4 \rceil)\mathrm{Re}(\zeta) + (\lceil u_2 \rceil + \lceil u_3 \rceil)\mathrm{Re}(\zeta^2)$. Obviously $\lceil u_1 \rceil + \lceil u_4 \rceil = \lceil u_1 + u_4 \rceil + q$, where q is either 0 or 1, and similarly $\lceil u_2 \rceil + \lceil u_3 \rceil = \lceil u_2 + u_3 \rceil + r$, with r either zero or 1. With the abbreviation

$$A = \lceil (k - \gamma_0)\tau + \gamma_1 + \gamma_4 \rceil \tau/2 - \lceil -(k - \gamma_0)\tau^{-1} + \gamma_2 + \gamma_3 \rceil \tau^{-1}/2$$

this leads to

$$\mathrm{Re}(M(t)) = k + A + q\mathrm{Re}(\zeta) + r\mathrm{Re}(\zeta^2).$$

It gives the four different horizontal coordinates of the left end-points of the horizontal edges in the stack. The right end-points are obtained by adding 1. So all the vertices involved here are lying on eight vertical lines. The figure formed by those eight lines has the central bar as an axis of symmetry, and that is why the name "central" was chosen. Since $\mathrm{Re}(\zeta) + \mathrm{Re}(\zeta^2) = -\frac{1}{2}$ the average of the eight horizontal coordinates is $k + A + \frac{1}{4}$, whence the central bar can be represented as $\mathrm{Re}(z) = k + A + \frac{1}{4}$.

From the arrowing of the horizontal vertices in the stack it is easy to derive that this central bar cuts the doubly arrowed horizontal edges at a point $\frac{1}{4}$ from the end, exactly in accordance with Figure 5.3.

These calculations did not yet use the condition that $\sum \gamma_j = 0$. Tilings without that condition were mentioned in [3] in connection with a riffle shuffle card trick, and it is actually the riffle shuffle arrangement in a stack that guarantees that there are only eight different horizontal coordinates.

But $\sum \gamma_j = 0$ gives a simplification:

$$-(k - \gamma_0)\tau^{-1} + \gamma_2 + \gamma_3 = \mu_0\tau - k\tau - k,$$

and since non-singularity is assumed, this cannot be an integer. If x is not an integer one has $-\lceil -x \rceil = \lceil x \rceil - 1$, which leads to the following formula for the central bar:

$$\mathrm{Re}(z) = (1 + \frac{1}{2}\tau^{-1})(k + \tau\lceil (k - \mu_0)\tau \rceil - \frac{1}{2}\tau).$$

For the singular cases the result (6.3.1) can now be derived by obvious limit operations.

6.5. The topological equivalence of pentagrid and central grid will now be established directly, on the basis of (6.3.1). Figure 6.1 presents an illustration. It looks reasonable to enlarge the pentagrid by a factor 5/2 when comparing it to the central grid (it is the same factor as in [1], section 5, formula above (5.3)), but for the topology of the grids such factors make no difference at all.

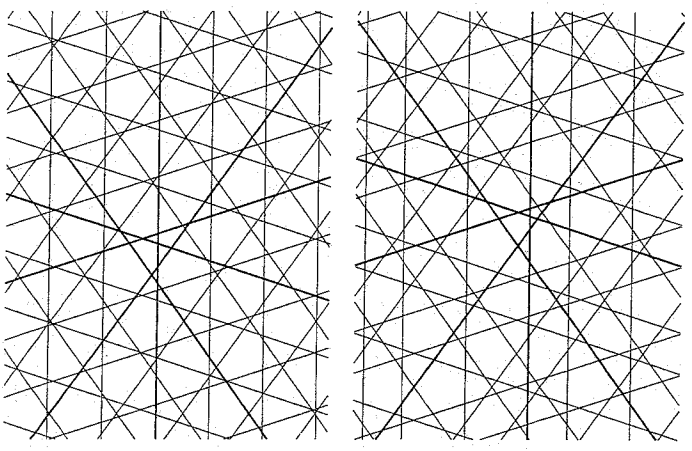

Fig. 6.1. On the left there is a piece of the pentagrid with parameters $\gamma_0 = 0.2$, $\gamma_1 = 0.4$, $\gamma_2 = 0.3$, $\gamma_3 = -0.8$, $\gamma_4 = -0.1$, on the right a corresponding piece of the central grid with the same parameters. The scale of the picture on the left is 5/2 times the one on the right. In order to facilitate the comparison of the topologies, the lines with $k_j = 0$ in the j-th subgrid are thicker than the others.

The method is as follows. Consider two grids of the form of section 6.1, given as $\mathrm{Re}(z\zeta^{-j}) = \alpha_{j,k}$ and $\mathrm{Re}(z\zeta^{-j}) = \beta_{j,k}$. It is assumed that for each j the $\alpha_{j,k}$ and the $\beta_{j,k}$ increase with k. And it is assumed that nowhere in the grids three lines pass through a point. In the topological correspondence the k-th line of the j-th subgrid of the α-grid will correspond to the k-th line of the j-th subgrid of the β-grid, for all j and k.

If p, q, r are straight lines forming a triangle, then they determine an orientation: seen from the inside of the triangle the circular order p, q, r is either clockwise or counter-clockwise. This gives the criterion for the topological equivalence in the grid: for any three lines in the α-grid the orientation has to be the same as for the corresponding three lines in the β-grid.

There are two kinds of triangles here. One is of the kind formed with values $j, j-1, j+1$, the other one with $j, j+2, j+3$ (j taken mod 5). The orientation is a simple matter of determinants. The result is as follows. The topological equivalence of the two grids is guaranteed if for all j and for all integers p, q, r the combination $\alpha_{j+1,p} + \alpha_{j-1,q} - \tau\alpha_{j,r}$ has the same sign as $\beta_{j+1,p} + \beta_{j-1,q} - \tau\beta_{j,r}$, and $\alpha_{j+2,p} + \alpha_{j-2,q} + \tau^{-1}\alpha_{j,r}$ has the same sign as $\beta_{j+2,p} + \beta_{j-2,q} + \tau^{-1}\beta_{j,r}$.

Since no three lines pass through a point none of these expressions is zero.

It is this very explicit condition that has to be verified for the regular pentagrid and the corresponding central grid. It will be established in sections 6.7-6.9 as a theorem on inequalities that can be understood independently of the previous sections. In order to get rid of the factor $(1 + \frac{1}{2}\tau^{-1})$ those sections work with θ instead of β, where $\beta = (1 + \frac{1}{2}\tau^{-1})\theta$. And for simplicity, sections 6.7-6.9 restrict themselves to $j = 0$; the other cases are completely similar.

6.6. If the conditions mentioned in section 6.5 are satisfied for α and β then they are obviously also satisfied for α and $\omega(\lambda)$, where $0 \leq \lambda \leq 1$, and $\omega(\lambda)$ is obtained by linear interpolation:

$$\omega(\lambda)_{j,k} = (1 - \lambda)\alpha_{j,k} + \lambda\beta_{j,k}.$$

This means that the the α-grid can be deformed continuously into the β-grid without ever changing the topology. It is even possible to push the λ beyond 1. The lower bound $\frac{1}{2}\tau$ in theorem 6.1 (section 6.7) has the effect that the topology of the $\omega(\lambda)$-grid remains the same over the interval $0 \leq \lambda < (1 - \frac{1}{2}(1 + \frac{1}{2}\tau^{-1})\tau)^{-1}$.

Section 4.5 discussed infinitesimal perturbations for the interpretation of the dual of a singular pentagrid. The same thing can now be achieved with the $\omega(\lambda)$, with small positive but not infinitesimal values of λ.

6.7. The sections 6.7-6.9 do not make use of anything said in the previous sections.

Let $\gamma_0, \ldots, \gamma_4$ be real numbers with $\gamma_0 + \ldots + \gamma_4 = 0$. For $j = 0, \ldots, 4$, $k_j \in \mathcal{Z}$, the real numbers $\alpha_{j,k}$ and $\theta_{j,k}$ are defined by

$$\alpha_{j,k} = k - \gamma_j, \quad \theta_{j,k} = k + \tau\lceil(k - \mu_j)\tau\rceil - \frac{1}{2}\tau,$$

where $\mu_j = \gamma_j - (\gamma_{j-1} + \gamma_{j+1})\tau^{-1}$, $\gamma_5 = \gamma_0$ and $\tau = \frac{1}{2}(-1 + \sqrt{5})$.

Let $k_0, \ldots, k_4$ be integers, and abbreviate

$$H = \alpha_{1,k_1} + \alpha_{4,k_4} - \tau\alpha_{0,k_0}, \quad L = \theta_{1,k_1} + \theta_{4,k_4} - \tau\theta_{0,k_0},$$

$$K = \alpha_{2,k_2} + \alpha_{3,k_3} + \tau^{-1}\alpha_{0,k_0}, \quad M = \theta_{2,k_2} + \theta_{3,k_3} + \tau^{-1}\theta_{0,k_0}.$$

With these abbreviations the following theorem can be proved:

Theorem 6.1. *If $H \neq 0$ then $L/H \geq \frac{1}{2}\tau$. If $K \neq 0$ then $M/K \geq \frac{1}{2}\tau$.*

6.8. Here is the proof of the first part of the theorem. H and L can be expressed like this:

$$H = k_1 + k_4 - \tau k_0 + \tau\mu_0,$$

$$L = k_1 + k_4 - \tau k_0 + \tau\lceil(k_1 - \mu_1)\tau\rceil + \tau\lceil(k_4 - \mu_4)\tau\rceil - \tau^2\lceil(k_0 - \mu_0)\tau\rceil - \tau + \tfrac{1}{2}\tau^2.$$

If x and y are real numbers then $\lceil x\rceil + \lceil y\rceil - \lceil x + y\rceil$ is either 0 or 1. So

$$\lceil(k_1 - \mu_1)\tau\rceil + \lceil(k_4 - \mu_4)\tau\rceil = \lceil(k_1 + k_4)\tau - (\mu_1 + \mu_4)\tau\rceil + p$$

with $p = 0$ or $p = 1$. Moreover

$$(k_0 - \mu_0)\tau = k_1 + k_4 - H,$$

and since $(\mu_1 + \mu_4)\tau = -\mu_0$,

$$(k_1 + k_4)\tau - (\mu_1 + \mu_4)\tau = H\tau^{-1} + k_0 - k_1 - k_4.$$

So there is a simple expression for L in terms of H and p:

$$L = \tau\lceil H\tau^{-1}\rceil - \tau^2\lceil -H\rceil - \tau + \frac{1}{2}\tau^2 + p\tau.$$

Now first assume $H > 0$. Use $p \geq 0$, $-\lceil -H\rceil \geq 0$, and note that if $0 < c < 1$ then $\lceil x\rceil \geq 1 - c + cx$ for all $x > 0$. The special case $x = H\tau^{-1}$, $c = \frac{1}{2}\tau$ leads to $L/H \geq \frac{1}{2}\tau$.

Next assume $H < 0$. Use $p \leq 1$, $\lceil -H\rceil \geq 1$, and note that if $0 < c < 1$ then $\lceil x\rceil \leq cx + c$ for all $x < 0$. With $x = H\tau^{-1}$, $c = \frac{1}{2}\tau$ it follows that $L/H \geq \frac{1}{2}\tau$.

6.9. The proof of the second part of the theorem is similar. K and M can be expressed as follows: $K = k_2 + k_3 + k_0\tau^{-1} - \tau\mu_0$ and

$$M = k_2 + k_3 + k_0\tau^{-1} + \tau\lceil(k_2 - \mu_2)\tau\rceil + \tau\lceil(k_3 - \mu_3)\tau\rceil + \lceil(k_0 - \mu_0)\tau\rceil - \tau - \frac{1}{2}.$$

Note that

$$\lceil(k_2 - \mu_2)\tau\rceil + \lceil(k_3 - \mu_3)\tau\rceil = \lceil(k_2 + k_3)\tau - (\mu_2 + \mu_3)\tau\rceil + q$$

where q is either 0 or 1. Moreover $(k_0 - \mu_0)\tau = K - k_0 - k_2 - k_3$, and since $\mu_2 + \mu_3 = \mu_0\tau$,

$$(k_2 + k_3)\tau - (\mu_2 + \mu_3)\tau = K\tau - k_0.$$

So M can be expressed in terms of K and q:

$$M = \tau\lceil K\tau\rceil + \lceil K\rceil + \tau q - \tau - \frac{1}{2}.$$

If $K > 0$ it can be used that $\lceil K\rceil \geq 1$, whence $M \geq \tau\lceil K\tau\rceil - \tau + \frac{1}{2}$. If $0 < c < 1$ then $\lceil x\rceil \geq 1 - c + cx$ for all $x > 0$. With $x = K\tau$, $c = \frac{1}{2}\tau^{-1}$ this leads to $M/K \geq \frac{1}{2}\tau$. If $K < 0$ one can use the inequalities $q \geq 0$, $\lceil K\rceil \leq 0$ and $\lceil x\rceil \leq cx + c$ ($x < 0$, $0 < c < 1$) with $x = K\tau$, $c = \frac{1}{2}\tau^{-1}$. This again leads to $M/K \geq \frac{1}{2}\tau$, and that finishes the proof of theorem 6.1.

References

1. N. G. de Bruijn, Algebraic theory of Penrose's non-periodic tilings of the plane. Kon. Nederl. Akad. Wetensch. Proc. Ser. A 84 (=Indagationes Mathematicae 43), 38-52 and 53-66 (1981). Reprinted in: P. J. Steinhardt and Stellan Ostlund: The Physics of Quasicrystals, World Scientific Publ., Singapore, New Jersey, Hong Kong.
2. N.G. de Bruijn, Dualization of multigrids. In: Proceedings of the International Workshop Aperiodic Crystals, Les Houches 1986. Journal de Physique, Vol.47, Colloque C3, supplement to nr. 7, July 1986, pp. 9-18.
3. N. G. de Bruijn, A riffle shuffle card trick and its relation to quasicrystal theory. Nieuw Archief Wiskunde (4) 5 (1987) 285-301.
4. N. G. de Bruijn, Symmetry and quasisymmetry. In: Symmetrie in Geistes- und Naturwissenschaft. Herausg. R. Wille. Springer Verlag 1988, pp. 215-233.
5. N. G. de Bruijn, Updown generation of Penrose tilings. Indagationes Mathematicae, N.S., 1, pp. 201-219 (1990).
6. Martin Gardner, Mathematical games. Extraordinary nonperiodic tiling that enriches the theory of tiles. Scientific American 236 (1) 110-121 (Jan. 1977).
7. Branko Grünbaum and G.C. Shephard. Tilings and patterns. New York, W.H. Freeman and Co. 1986.
8. R. Penrose. Pentaplexity. Mathematical Intelligencer vol 2 (1) pp. 32-37 (1979).
9. J. E. S. Socolar and P. J. Steinhardt. Quasicrystals. II. Unit cell configurations. Physical Rev. B Vol. 34 (1986), 617-647. Reprinted in: P.J. Steinhardt and Stellan Ostlund: The Physics of Quasicrystals, World Scientific Publ., Singapore, New Jersey, Hong Kong.