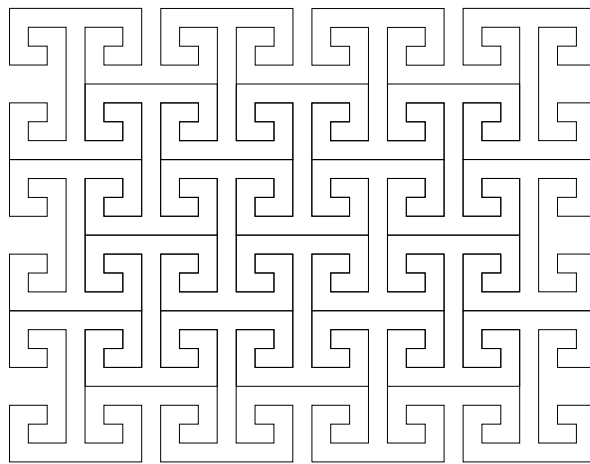


Tilings and Patterns



Fall 2025

by Jarkko Kari, University of Turku

Contents

1	Introduction	1
2	Symmetries	1
2.1	Isometries of the plane	1
2.2	Fixed points	5
2.3	Symmetries of a set of points	7
2.4	Products of two reflections	8
2.5	Parity	10
2.6	Odd isometries	12
2.7	Rosette groups	13
2.8	Conjugacy	16
2.9	Frieze groups	18
2.10	Wallpaper groups	21
2.11	Final remarks on discrete symmetry groups	36
3	Tilings	37
3.1	Basic definitions	38
3.2	Tilings by regular polygons	42
4	Wang tiles	48
4.1	Periodic tilings	49
4.2	Compactness principle	51
4.3	Robinson's aperiodic tile set	52
4.4	An aperiodic set of 14 Wang tiles	56
5	Undecidable problems concerning tiles	61
5.1	Turing machines	63
5.2	The tiling problem with a seed tile	65
5.3	Finite systems of forbidden patterns	67
5.4	The periodic tiling problem	69
5.5	The tiling problem	73
5.6	The completion problem	78
5.7	Beyond aperiodicity: arecursive tile sets	81
6	Compact topology on Wang tilings	84
6.1	Review of topology and metric spaces	85
6.2	Basic facts about the configuration space	91
6.3	Subshifts	92
6.4	Orbits, transitivity and minimality	93
6.5	Periodicity and recurrence properties	95
6.6	Equicontinuity and isolated points	98
7	A brief revisit to tilings by polygons	99
7.1	Substitutions	101
7.2	Amman's aperiodic tile set	103
7.3	The extension and the periodicity theorems	107
7.4	Hat: an aperiodic monotile	108
7.5	Open problems	119

1 Introduction

Informally, a tiling is a covering of the plane with tiles of various shapes in such a way that the tiles do not overlap each other. Often the tiles have simple shapes (e.g. polygons), and typically only a small number of different shapes are used in each tiling. Such tilings are everywhere around us: in pavements, quilt patterns, fabrics, brick walls, carpets, etc. Interest to decorative tilings is very old: Moors are an example of a culture that produced complex geometric patterns in tilings – famous examples can be found in the Alhambra at Granada, Spain.

In this course we learn about mathematical concepts relevant to tilings and patterns. The mathematical tools we use include high-school level geometry, elementary group theory, some topology, combinatorics and computation theory. After initial geometric considerations we work in detail on some computational questions on tilings, including decidability aspects. The basics of computation theory and other required material are provided during the course as needed, so that the course is made as self-contained as possible. In some instances we may rely on theorems from other fields that are presented without proofs, and in these instances an interested reader is directed to literature or other courses offered on these topics for more details and precise proofs.

2 Symmetries

Let us begin by investigating the fundamental concepts of symmetry.

2.1 Isometries of the plane

A plane isometry is any function $\alpha : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ that preserves distance:

$$\forall (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2 : d(\alpha(x_1, y_1), \alpha(x_2, y_2)) = d((x_1, y_1), (x_2, y_2))$$

where the distance $d : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ is the usual Euclidean distance defined by

$$d((x_1, y_1), (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

In other words, α moves the points of the plane in a "rigid" motion that does not change any distances.

In these notes we'll denote points of the plane by capital letters, so the isometry property will be written as

$$\forall P, Q \in \mathbb{R}^2 : d(\alpha(P), \alpha(Q)) = d(P, Q).$$

Our first theorem states that an isometry is necessarily a bijection (that is, both one-to-one and onto). This implies that it has an inverse function. This inverse function is also an isometry.

Theorem 2.1 *An isometry is a bijection. Its inverse function is an isometry.*

Proof. Let α be an isometry. It is trivial that α is one-to-one (also the term "injective" is used). Namely, if $\alpha(P) = \alpha(Q)$ then

$$d(P, Q) = d(\alpha(P), \alpha(Q)) = 0,$$

which means that $P = Q$.

The proof that α is onto (also the term "surjective" is used) is more difficult, and is therefore left as a homework problem ;-)

Let $P, Q \in \mathbb{R}^2$ be arbitrary and denote $P' = \alpha^{-1}(P)$ and $Q' = \alpha^{-1}(Q)$. Then $P = \alpha(P')$ and $Q = \alpha(Q')$ so $d(P', Q') = d(P, Q)$, which proves that the inverse function α^{-1} preserves distance.

□

Our next observation states that isometries preserve shapes. More precisely, let us show that an isometry maps every line into a line, every triangle into a triangle (we say that it preserves lines and triangles), and the angle between two lines remains the same. Also betweenness and midpoints are preserved.

Theorem 2.2 *An isometry preserves lines, triangles, betweenness, midpoints, sizes of angles, and perpendicularity and parallelism of lines.*

Proof. Let α be an isometry. Let us prove the preservation of

- betweenness and midpoints: If three points P, Q and R are collinear, with point R between points P and Q , then $d(P, R) + d(R, Q) = d(P, Q)$. But then we have also

$$d(P', R') + d(R', Q') = d(P', Q')$$

where $P' = \alpha(P)$, $Q' = \alpha(Q)$ and $R' = \alpha(R)$. This means that points P', Q' and R' are also collinear, with R' between points P' and Q' . So betweenness is preserved.

Since the inverse α^{-1} is also an isometry, the preservation works also in the inverse direction. In other words, R is between P and Q if and only if $\alpha(R)$ is between $\alpha(P)$ and $\alpha(Q)$.

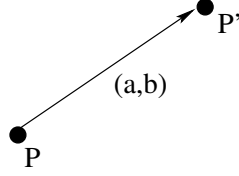
In the special case that R is the midpoint between P and Q we have that $d(P, R) = d(R, Q)$, so also $d(P', R') = d(R', Q')$, which means that R' is the midpoint between P' and Q' .

- triangles: Let $\triangle ABC$ be a triangle and, as usual, let us denote $A' = \alpha(A)$, $B' = \alpha(B)$ and $C' = \alpha(C)$. The triangle consists of those points P that are between A and B , between A and C , or between B and C . This is equivalent to $P' = \alpha(P)$ being between A' and B' , between A' and C' , or between B' and C' . Hence the image of triangle $\triangle ABC$ is the triangle $\triangle A'B'C'$.
- lines: Let m be a line, and let A and B be two points on the line. Then the line consists exactly of those points P such that (i) P is between A and B , (ii) A is between B and P , or (iii) B is between A and P . This is equivalent to $P' = \alpha(P)$ being such that (i) P' is between A' and B' , (ii) A' is between B' and P' , or (iii) B' is between A' and P' , where $A' = \alpha(A)$ and $B' = \alpha(B)$, which is equivalent to P' being on the line through points A' and B' .
- parallelism and perpendicularity of lines, as well as angles between lines: Take two different lines l and m . If they are parallel then they have no common points. Because α is one-to-one, their images $\alpha(l)$ and $\alpha(m)$ do not have any common points either, so they are parallel lines. Assume then that l and m are not parallel, in which case they intersect in one point P at some angle Θ . Let A and B be points of the lines l and m such that angle APB is of size Θ . Then the triangle $\triangle APB$ is congruent with its image $\triangle A'P'B'$ as the two triangles have same sides (SSS). Therefore the angle $A'P'B'$ is the same as the angle APB . In particular, l and m are perpendicular if and only if the angle is 90° , so also perpendicularity is preserved.

□

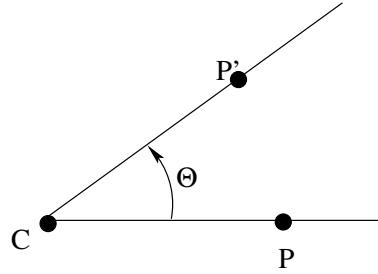
The trivial isometry is the identity function ι that does not move any points: $\iota(P) = P$ for all $P \in \mathbb{R}^2$. Let us look into some non-trivial examples of isometries.

Example 1. Let $A = (a, b) \in \mathbb{R}^2$. A translation by vector $A = (a, b)$ shifts every point (x, y) into position $(x + a, y + b)$. We denote a translation by vector A as τ_A .



Every translation is clearly an isometry. Trivial translation $\tau_{(0,0)}$ is the trivial isometry ι . □

Example 2. Let $C \in \mathbb{R}^2$ be a point, and $\Theta \in \mathbb{R}$ an angle. The rotation $\rho_{C,\Theta}$ by the (directed) angle Θ about C is the isometry that fixes point C , and otherwise takes point $P \neq C$ into the point P' where $d(C, P) = d(C, P')$ and Θ is the directed angle from CP to CP' :

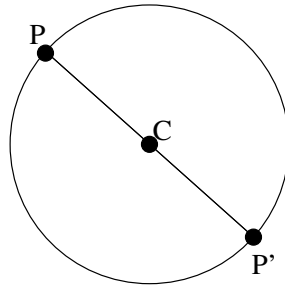


In terms of analytic geometry we say that point (x, y) is mapped to point (x', y') where

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos \Theta & -\sin \Theta \\ \sin \Theta & \cos \Theta \end{pmatrix} \begin{pmatrix} x - c_x \\ y - c_y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix}$$

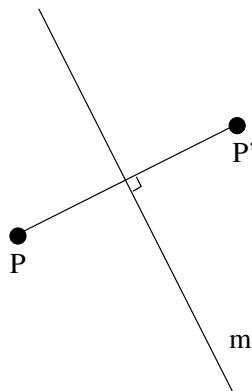
where $C = (c_x, c_y)$. Point C is called the center of the rotation. The trivial rotation $\rho_{C,0}$ by the angle 0° is the trivial isometry ι .

If $\Theta = 180^\circ$ we get a special case of the rotation called the halfturn about point C , or the reflection in point C . Every point P is mapped to the point P' such that the center C is the midpoint between P and P' :



Because halfturn about point C is an important particular case, we sometimes denote it by the special symbol σ_C . □

Example 3. Let m be a line. The reflection σ_m in line m is the mapping that does not move the points of line m , but any point P outside line m is moved to the point P' such that line m is the perpendicular bisector of segment PP' .



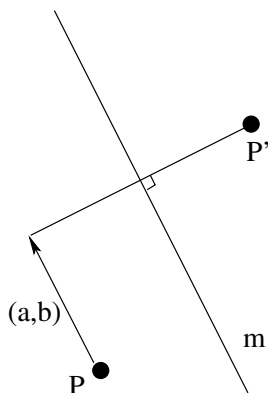
It follows immediately from the definition that $\sigma_m^{-1} = \sigma_m$, that is, the reflection σ_m is its own inverse. Isometries that are their own inverses are called involutions.

□

Example 4. A Glide reflection is a composition of a translation and a reflection in line m that is parallel with the direction of the translation. Let $A = (a, b) \in \mathbb{R}^2$ a vector of translation, and let m be a line parallel to A , that is,

$$m = \{(c, d) + t(a, b) \mid t \in \mathbb{R}\}$$

where (c, d) is some point of the line. The glide reflection $\gamma_{m, (a, b)}$ they specify reflects the points in line m and then translates them by vector A . In this particular case it does not matter in which order the two operations are performed: we may as well translate first and reflect later.



Line m is called the axis of the glide reflection. Notice that glide reflections with trivial translation vectors $A = (0, 0)$ are exactly the reflections.

□

Later we'll see that our four examples exhaust all possibilities: translations, rotations, reflections and glide reflections are the only isometries of the plane. (In fact, since reflection is a special type of glide reflection we can say that all isometries are translations, rotations or glide reflections.)

The composition $\alpha \circ \beta$ of two functions α and β is the function that first applies β to a point, and then applies α to the result, that is,

$$(\alpha \circ \beta)(x) = \alpha(\beta(x)).$$

If α and β are isometries then also their composition $\alpha \circ \beta$ is an isometry. Indeed, for any two points P and Q we have

$$d((\alpha \circ \beta)(P), (\alpha \circ \beta)(Q)) = d(\alpha(\beta(P)), \alpha(\beta(Q))) = d(\beta(P), \beta(Q)) = d(P, Q).$$

Function composition $\circ$ is an associative operation, and since the identity function ι and the inverses of all isometries are also isometries, we have the following theorem:

Theorem 2.3 *The set of plane isometries forms a group $\mathcal{I}$ under the operation of composition.*

□

We frequently drop the group operation sign " $\circ$ " and simply write $\alpha\beta$ for $\alpha \circ \beta$. We then say that $\alpha\beta$ is the product of operations α and β . We also do not need to use parentheses in products as, because of the associativity, $\alpha(\beta\gamma) = (\alpha\beta)\gamma$. We simply write this as $\alpha\beta\gamma$. However, remember that the group of isometries is not commutative (=abelian) as in most cases $\alpha\beta \neq \beta\alpha$.

An element $\alpha \in \mathcal{I}$ is called an involution if $\alpha^2 = \iota$. Examples of involutions include all reflections in lines, as well as all halfturns. In fact, no other involutions exist. Review the following terms of group theory:

- generator set (=set of group elements such that every element of the group is a product of generators and their inverses),
- cyclic group (=a group that is generated by one element)
- order of a group (=number of elements. If the group contains an infinite number of elements then the group is called infinite, otherwise it is finite.)
- subgroup (=a subset of the group that is closed under the group operation and the operation of taking the inverse element. A subgroup itself is a group under the same group operation)
- cancellation laws ($\alpha\beta = \alpha\gamma$ implies $\beta = \gamma$, and $\beta\alpha = \gamma\alpha$ implies $\beta = \gamma$.)

In the rest of this chapter we try to understand the structure of the group $\mathcal{I}$. We want to show that our examples exhaust all possibilities, and to find out how the group operation combines these isometries.

2.2 Fixed points

The two main results of this section are the following:

1. To verify that two given isometries α and β are the same, it is sufficient to verify that they agree on some three points that are not collinear (Corollary 2.6).
2. Every isometry is a product of at most three reflections (Corollary 2.7).

We say that P is a fixed point of isometry α if $\alpha(P) = P$. We also say that α fixes point P .

Lemma 2.4 *If an isometry α fixes two distinct points P and Q , then it fixes every point of the line m that contains P and Q .*

Proof. Assume that α fixes points P and Q of line m , and let R be any point of the line m . Because α preserves lines, $\alpha(R)$ is on the same line with $\alpha(P) = P$ and $\alpha(Q) = Q$, that is, $\alpha(R)$ is on line m . Because $d(\alpha(R), P) = d(R, P)$ and $d(\alpha(R), Q) = d(R, Q)$, we must have $\alpha(R) = R$. (There are two points at distance $d(R, P)$ from P , and these two points have different distances from point Q . So only one of these two points can have distance $d(R, Q)$ from Q , namely point R .)

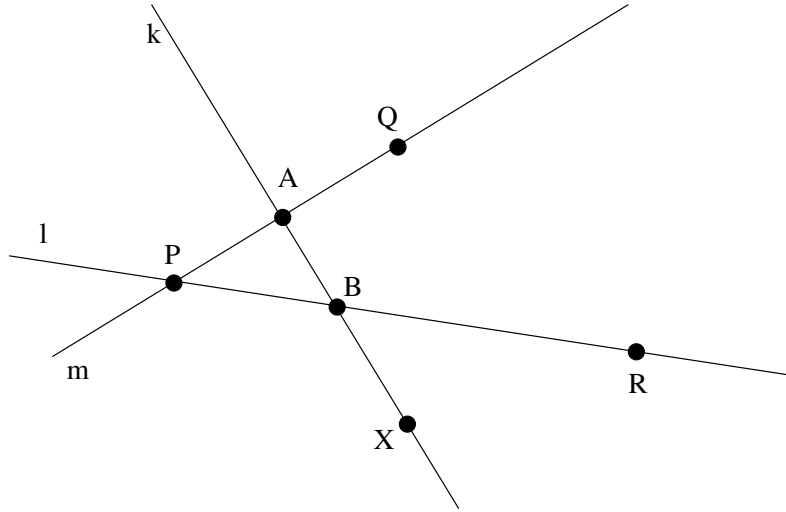
□

Consider three points P, Q and R that are not collinear, i.e. are not on the same line. As a corollary of the next theorem we get that their images $\alpha(P), \alpha(Q)$ and $\alpha(R)$ uniquely determines the isometry α . We also prove that every isometry is a product of at most three reflections.

Theorem 2.5 *Let α be an isometry.*

1. *If α fixes three non-collinear points, then $\alpha = \iota$.*
2. *If α fixes two points then $\alpha = \iota$ or α is a reflection.*
3. *If α fixes exactly one point then α is a product of two reflections.*

Proof. 1. Assume that α fixes three non-collinear points P, Q and R . Let m and l be the lines that contain P and Q , and P and R , respectively. According to Lemma 2.4, α fixes all points that belong to lines m or l . Let X be an arbitrary point outside lines m and l . There exists a line k that goes through X and intersects m and l at distinct points A and B .



Because α fixes A and B then, according to Lemma 2.4, it also fixes all points of line k , which means that it also fixes point X . As X was an arbitrary point, we conclude that α fixes all points of the plane, so $\alpha = \iota$.

2. Assume then that α fixes two distinct points P and Q , and suppose that $\alpha \neq \iota$. Then there exists some point R such that $\alpha(R) \neq R$. Notice that P, Q and R cannot be collinear (Lemma 2.4). Denote $R' = \alpha(R)$, and let m be the perpendicular bisector of the segment RR' . Then $R' = \sigma_m(R)$ where σ_m is the reflection in line m . Because $d(R', P) = d(R, P)$ and $d(R', Q) = d(R, Q)$, points P and Q are on the perpendicular bisector m . We have $\sigma_m(P) = P$ and $\sigma_m(Q) = Q$. The isometry $\sigma_m^{-1}\alpha$ hence fixes three non-collinear points P, Q and R so, according to case 1 of the theorem, $\sigma_m^{-1}\alpha = \iota$. This proves that $\alpha = \sigma_m$ is a reflection.

3. Assume that isometry α fixes exactly one point P . Let Q a different point, so $Q' = \alpha(Q)$ is different from Q . Let l be the perpendicular bisector of the segment QQ' . Triangle $\triangle QPQ'$ is isosceles, so the point P is on the line l . Then $\sigma_l^{-1}\alpha$ fixes two points P and Q , so according to case 2 either $\sigma_l^{-1}\alpha = \iota$ or $\sigma_l^{-1}\alpha = \sigma_m$ for some line m . The first alternative $\alpha = \sigma_l$ is not possible because then α fixes more than one point – it fixes all points of line l . So we must have the second alternative $\alpha = \sigma_l\sigma_m$.

□

Corollary 2.6 *If α and β are two isometries such that $\alpha(P) = \beta(P)$, $\alpha(Q) = \beta(Q)$ and $\alpha(R) = \beta(R)$, and points P, Q and R are not collinear, then $\alpha = \beta$.*

Proof. Isometry $\alpha^{-1}\beta$ fixes non-collinear points P, Q and R , so $\alpha^{-1}\beta = \iota$. This implies $\alpha = \beta$. □

Corollary 2.7 *Every isometry is a product of at most three reflections.*

Proof. If α fixes at least one point then, according to the theorem, α is a product of at most two reflections. Assume then that α does not fix any points. Let P be an arbitrary point, and let m be the perpendicular bisector of the segment $P\alpha(P)$. Then $\sigma_m^{-1}\alpha$ fixes point P , so $\sigma_m^{-1}\alpha$ is a product of at most two reflections and, therefore, α is a product of at most three reflections. □

The proofs provide a simple method of finding the reflections when we know the images $P_0 = \alpha(P)$, $Q_0 = \alpha(Q)$ and $R_0 = \alpha(R)$ of three given non-collinear points P, Q and R . We simply find reflections that match the points one-by-one:

1. If $P \neq P_0$ then we first reflect in line m that is the perpendicular bisector of the segment PP_0 . This maps P to its correct position P_0 . Let Q' and R' be the images of Q and R under the first reflection.
2. If $Q' \neq Q_0$ then we reflect in line l that is the perpendicular bisector of the segment $Q'Q_0$. Notice that point P_0 is on this bisector because $d(P_0, Q_0) = d(P, Q) = d(P_0, Q')$. After the second reflection, points P and Q have been mapped to their correct positions P_0 and Q_0 . Let R'' be the image of R after the first two reflections.
3. If $R'' \neq R_0$ then we finally reflect in line k that is the perpendicular bisector of R'' and R_0 . It is easy to see that P_0 and Q_0 are on this bisector:

$$d(P_0, R_0) = d(P, R) = d(P_0, R'') \quad \text{and} \quad d(Q_0, R_0) = d(Q, R) = d(Q_0, R'').$$

After steps 1–3, points P, Q and R have been mapped in their correct positions P_0, Q_0 and R_0

2.3 Symmetries of a set of points

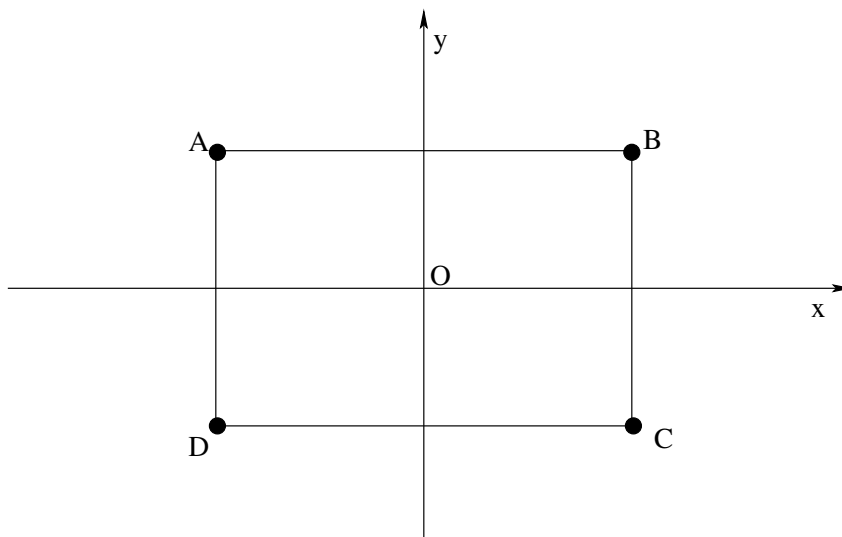
Let $s \subseteq \mathbb{R}^2$ be a set of points. We say that isometry α is a symmetry of set s iff $\alpha(s) = s$.

Theorem 2.8 *Let $s \subseteq \mathbb{R}^2$ be arbitrary. The symmetries of s form a subgroup of $\mathcal{I}$, the group of isometries.*

Proof. Every set has at least one symmetry, namely the trivial isometry ι . If $\alpha(s) = s$ then $\alpha^{-1}(s) = \alpha^{-1}(\alpha(s)) = s$, so the inverse of each symmetry of s is also a symmetry of s . Let α and β be two symmetries of s . Then $\alpha\beta(s) = \alpha(s) = s$ so the product $\alpha\beta$ is also a symmetry of s . □

The set of symmetries of s is called the symmetry group of s . Notice that $\mathcal{I}$ itself is the symmetry group of $s = \mathbb{R}^2$.

Example 5. Let s be a rectangle $ABCD$ that is not a square. Let us position s in such a way that its center is at the origin $(0, 0)$, and its sides are parallel to the x - and y -axes.



Any symmetry of s must permute the corners of the rectangle. Corner A may be mapped into any of the four corners A, B, C and D , after which the images of the other corners B, C and D are uniquely determined. We proved in the previous section that three non-collinear points A, B and C determine the entire isometry (Corollary 2.6), so the symmetry group s contains exactly four symmetries. These are ι , two reflections σ_h and σ_v in the x - and y -axes, and the halfturn σ_O about the origin O . These form Klein's Vierergruppe V_4 .

□

Example 6. If s is a square $ABCD$ then its symmetry group contains eight elements, so a square is "more" symmetric than a non-square rectangle. In the square we may map the corner A into any of the four corners, after which corner B has still two possible images. Then the images of C and D are uniquely determined.

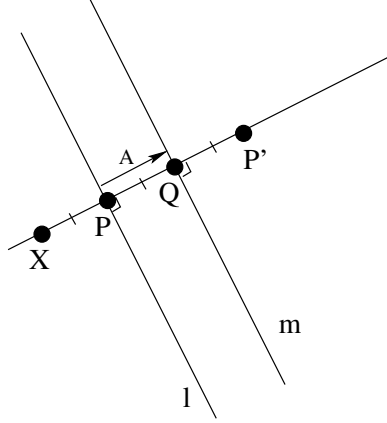
□

2.4 Products of two reflections

We know that every isometry is a product of at most three reflections. In order to characterize all isometries we need to investigate the products of two or three reflections. Let us start by products of two reflections.

Theorem 2.9 *The product of two reflections in parallel lines m and l is a translation in the direction perpendicular to l and m by a distance that is twice the distance from l to m . Conversely, every translation is a product of two reflections in parallel lines, both perpendicular to the direction of the translation. One of the lines can be chosen freely (as long as it is perpendicular to the translation).*

Proof. Let m and l be two parallel lines. If $m = l$ then $\sigma_m \sigma_l = \iota = \tau_{(0,0)}$. Assume then that $m \neq l$. Let A be the vector from l to m that is perpendicular to m and l . To prove that $\sigma_m \sigma_l = \tau_{2A}$ it is enough to show that $\sigma_m \sigma_l(P) = \tau_{2A}(P)$ for every point P of line l , and that $\sigma_m \sigma_l(X) = \tau_{2A}(X)$ for some point X outside of line l . Then the result follows from Corollary 2.6.



Referring to the figure above, we have that for every $P \in l$

$$\sigma_m \sigma_l(P) = \sigma_m(P) = P' = \tau_{2A}(P).$$

Analogously, by reversing the roles of lines m and l , we have that for an arbitrary $Q \in m$

$$\sigma_l \sigma_m(Q) = \tau_{-2A}(Q).$$

Let $X = \sigma_l \sigma_m(Q) = \tau_{-2A}(Q)$. Then X is not on line l , and

$$\sigma_m \sigma_l(X) = \sigma_m \sigma_l \sigma_l \sigma_m(Q) = Q = \tau_{2A} \tau_{-2A}(Q) = \tau_{2A}(X).$$

The second part of the theorem follows directly from the first part: Let τ be a non-trivial translation, and let P be an arbitrary point and $P' = \tau(P)$. Let l and m be the lines perpendicular to the segment PP' through P and the midpoint of PP' , respectively. Then, according to the first part, $\sigma_m \sigma_l = \tau$. □

Corollary 2.10 *The product of three reflections in three parallel lines is a reflection in a parallel line.*

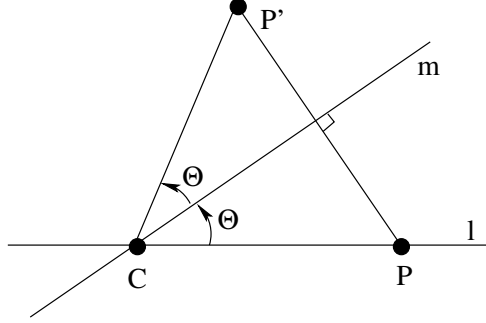
Proof. Let l, m and n be any three parallel lines. Let p be a fourth parallel line whose distance from line n is the same as the distance of line l from line m . Then $\sigma_l \sigma_m$ and $\sigma_p \sigma_n$ are the same translation. Multiplying by σ_n from the right gives $\sigma_l \sigma_m \sigma_n = \sigma_p$. □

Consider then two reflections in lines that are not parallel:

Theorem 2.11 *The product of two reflections in intersecting lines is a rotation about the point of intersection, and the angle of the rotation is twice the angle between the lines. Conversely, every rotation about point C is a product of two reflections in lines through point C . One of these lines can be chosen freely.*

Proof. Let l and m be lines that intersect at point C . Let Θ be the directed angle between them measured from l to m . Let us prove that $\sigma_m \sigma_l = \rho_{C, 2\Theta}$ by showing that $\sigma_m \sigma_l$ and $\rho_{C, 2\Theta}$ agree on three non-collinear points: all points of line l , and one point X that is outside of line l .

First, as all the isometries σ_m , σ_l and $\rho_{C, 2\Theta}$ fix point C , we have $\sigma_m \sigma_l(C) = C = \rho_{C, 2\Theta}(C)$. Let $P \neq C$ be a point on line l , and let $P' = \rho_{C, 2\Theta}(P)$. Line m is the perpendicular bisector of PP' , so $P' = \sigma_m(P) = \sigma_m \sigma_l(P)$.



So far we have proved that $\sigma_m \sigma_l(P) = \rho_{C, 2\Theta}(P)$ for all $P \in l$. Analogously, by reversing the roles of lines m and l , we have that $\sigma_l \sigma_m(Q) = \rho_{C, -2\Theta}(Q)$ for an arbitrary point $Q \neq C$ of line m . Denote $X = \sigma_l \sigma_m(Q) = \rho_{C, -2\Theta}(Q)$. Then X is not on line l and

$$\sigma_m \sigma_l(X) = \sigma_m \sigma_l \sigma_l \sigma_m(Q) = Q = \rho_{C, 2\Theta} \rho_{C, -2\Theta}(Q) = \rho_{C, 2\Theta}(X).$$

To prove the second part of the theorem, consider an arbitrary rotation $\rho_{C, \Theta}$. Let l be an arbitrary line through the center C of the rotation, and let m be the line through point C that meets line l in the directed angle $\Theta/2$. According to the first part of the theorem we have $\sigma_m \sigma_l = \rho_{C, \Theta}$.

□

Corollary 2.12 *Halfturn σ_C is the product of two reflections in any two perpendicular lines through C . In particular, reflections in perpendicular lines commute.*

□

Corollary 2.13 *The product of three reflections in lines through the common point C is a reflection in a line through point C .*

Proof. As in the proof of Corollary 2.10, let l, m and n be any three lines through point C . Let p be a fourth line through C that forms with line n the same angle as line l forms with line m . Then $\sigma_l \sigma_m$ and $\sigma_p \sigma_n$ are the same rotation about point C . Multiplying by σ_n from the right gives $\sigma_l \sigma_m \sigma_n = \sigma_p$.

□

2.5 Parity

As we proved previously, all isometries are products of some reflections, in fact, of at most three reflections. The representation of an isometry as a product of reflections is, however, not unique. For example, we can always add $\sigma_m \sigma_m$ to the end of any sequence of reflections, thus increasing the number of reflections in the sequence by two. However, it turns out that the parity of the number of reflections is always the same. We call isometry α even if it is a product of an even number of reflections, and odd if it is a product of an odd number of reflections. Next we want to show that no isometry can be both even and odd at the same time, that is, even and odd products of reflections can never be equal.

First we can make the following easy observation: A product of two reflections is not a reflection. Indeed, we know from the results of the previous section that a product of two reflections is either a translation or a rotation. Translations have no fixed points, rotations have exactly one fixed point, and the trivial isometry ι fixes all points. In contrast, the fixed points of a reflection form a line. So $\sigma_m \sigma_l \neq \sigma_k$ for all lines m, l and k .

The following theorem provides a method of reducing by two the number of terms in any long product of reflections:

Theorem 2.14 *A product of four reflections is a product of two reflections.*

Proof. We use the following lemma twice:

Lemma 2.15 *If m and l are two lines and P is a point, then there are lines p and q such that $\sigma_m\sigma_l = \sigma_p\sigma_q$, and line q contains point P .*

Proof of the lemma. If m and l are parallel, then we choose as q the line that is parallel to m and l and goes through point P . By corollary 2.10 we have $\sigma_m\sigma_l\sigma_q = \sigma_p$ for some line p , so $\sigma_m\sigma_l = \sigma_p\sigma_q$.

If m and l intersect at some point Q , then we choose as q a line through points P and Q . By corollary 2.13 we have $\sigma_m\sigma_l\sigma_q = \sigma_p$ for some line p , so $\sigma_m\sigma_l = \sigma_p\sigma_q$. □

Consider a product $\sigma_m\sigma_l\sigma_k\sigma_n$ of four reflections. Let P be an arbitrary point on line n . According to the lemma, $\sigma_l\sigma_k = \sigma_p\sigma_q$ where line q contains point P . Then we apply the lemma again: $\sigma_m\sigma_p = \sigma_r\sigma_s$ where s contains point P . We have

$$\sigma_m\sigma_l\sigma_k\sigma_n = \sigma_m\sigma_p\sigma_q\sigma_n = \sigma_r\sigma_s\sigma_q\sigma_n,$$

and lines n, q and s go through point P . Then, by Corollary 2.13 the product $\sigma_s\sigma_q\sigma_n = \sigma_t$ for some line t . Hence

$$\sigma_m\sigma_l\sigma_k\sigma_n = \sigma_r\sigma_t. \quad \square$$

Corollary 2.16 *A product of three reflections cannot equal a product of two reflections.*

Proof. Assume that

$$\sigma_m\sigma_l\sigma_k = \sigma_n\sigma_r.$$

Multiplying from left by σ_n gives

$$\sigma_n\sigma_m\sigma_l\sigma_k = \sigma_r.$$

According to the theorem there exist lines p and q such that

$$\sigma_n\sigma_m\sigma_l\sigma_k = \sigma_p\sigma_q,$$

so $\sigma_p\sigma_q = \sigma_r$, a contradiction. □

Corollary 2.17 *A product of an even number of reflections cannot equal a product of an odd number of reflections.*

Proof. By using the theorem we can reduce by two the number of reflections in any product of at least four reflections. In this way, any even length sequence can be reduced into a product of two reflections, and any odd length sequence reduces into a length one or a length three sequence. As a product of two reflections cannot equal a product of one or three reflections, we have the desired result. □

Now we know that every isometry is either even or odd, but not both. Notice that odd isometries correspond to "flipping" the plane over, turning all shapes into their mirror images. As every even isometry is a product of two reflections, we have

Theorem 2.18 *Even isometries are exactly the translations and the rotations.*

□

Notice also that even isometries form a subgroup of $\mathcal{I}$. Indeed, the inverse of the even isometry $\sigma_m\sigma_l$ is the even isometry $\sigma_l\sigma_m$, and the product of two even isometries $\sigma_m\sigma_l$ and $\sigma_n\sigma_p$ is the even isometry $\sigma_m\sigma_l\sigma_n\sigma_p$. Let us denote the group of even isometries by $\mathcal{E}$.

2.6 Odd isometries

Let's turn our attention to odd isometries. The goal of this section is to prove that every odd isometry is a glide reflection (where we understand that a plain reflection is also a glide reflection with a zero glide.) Recall that we use the notation σ_P for the halfturn about point P .

Lemma 2.19 *Isometry α is a glide reflection if and only if $\alpha = \sigma_P\sigma_l$ for some point P and line l . This is also equivalent to $\alpha = \sigma_k\sigma_Q$ for some line k and point Q .*

Proof. Let α be a glide reflection. By definition, $\alpha = \sigma_m\tau_A$ where the translation τ_A is in the direction of line m . By Theorem 2.9 $\tau_A = \sigma_k\sigma_l$ where lines k and l are perpendicular to line m . We have $\alpha = \sigma_m\sigma_k\sigma_l$. Corollary 2.12 states that the product $\sigma_m\sigma_k$ of two reflections in perpendicular lines is the halfturn σ_P about the intersection point P of lines m and k . We have

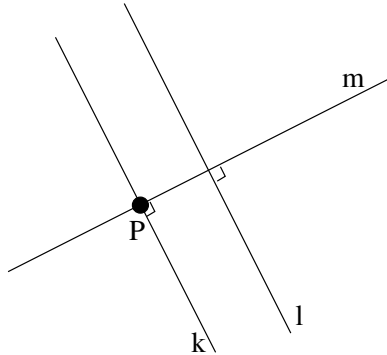
$$\alpha = \sigma_m\sigma_k\sigma_l = \sigma_P\sigma_l$$

as desired. We also have $\sigma_P = \sigma_k\sigma_m$, so

$$\alpha = \sigma_P\sigma_l = \sigma_k\sigma_m\sigma_l = \sigma_k\sigma_Q$$

where Q is the point where perpendicular lines m and l intersect.

For the converse claim, assume that α is the isometry $\sigma_P\sigma_l$ for some point P and line l . Let k be the line through point P that is parallel to line l , and let m be the line through point P that is perpendicular to lines k and l . Then, by Corollary 2.12, $\sigma_P = \sigma_m\sigma_k$.



We have

$$\alpha = \sigma_P\sigma_l = \sigma_m\sigma_k\sigma_l = \sigma_m\tau_A$$

where τ_A is in the direction of line m . Hence α is a glide reflection.

Analogously, if $\alpha = \sigma_k\sigma_Q$, and lines m and l go through point Q , and l is parallel and m perpendicular to k , then

$$\alpha = \sigma_k\sigma_Q = \sigma_k\sigma_m\sigma_l = \sigma_m\sigma_k\sigma_l = \sigma_m\tau_A$$

where A is in the direction of line m .

□

Now we are able to prove the main result on odd isometries:

Theorem 2.20 *Every odd isometry is a glide reflection.*

Proof. Let α be an odd isometry. Then it is either a reflection (which is a special type of a glide reflection) or a product of three reflections. Let $\alpha = \sigma_m \sigma_l \sigma_k$. Let P be an arbitrary point on line k . By Lemma 2.15 there exist lines p and q such that $\sigma_m \sigma_l = \sigma_p \sigma_q$ and line q goes through point P . We have

$$\alpha = \sigma_m \sigma_l \sigma_k = \sigma_p \sigma_q \sigma_k,$$

and $P \in k, q$. Let n be the line through point P that is perpendicular to line p . As lines n, q and k all go through point P , the product $\sigma_n \sigma_q \sigma_k$ is some reflection σ_r , see Corollary 2.13. Then $\sigma_q \sigma_k = \sigma_n \sigma_r$, and

$$\alpha = \sigma_p \sigma_n \sigma_r.$$

Lines n and p are perpendicular, so the product $\sigma_p \sigma_n$ is a halfturn σ_Q , where Q is the point where n and p intersect. We have

$$\alpha = \sigma_Q \sigma_r,$$

and it now follows from Lemma 2.19 that α is a glide reflection. □

Now we have classified all isometries of the plane. Even isometries are translations and rotations, and odd isometries are glide reflections (including reflections without glides).

2.7 Rosette groups

Rosette groups are the finite subgroups of $\mathcal{I}$. In this section we prove that the rosette groups are the cyclic groups C_n and the dihedral groups D_n , for $n \geq 1$, defined as follows:

The cyclic group C_n consists of n rotations about the same center P . It is generated by the single rotation $\rho = \rho_P, \frac{360^\circ}{n}$, so the elements of C_n are $\rho, \rho^2, \dots, \rho^n = \iota$. Notice that strictly speaking there are infinitely many groups C_n as the center P can be any point of the plane, but they are all obviously isomorphic with each other.

The dihedral group D_n includes C_n , and in addition it contains reflections in n lines that meet at P (the center of the rotations) at angles that are multiples of $\frac{360^\circ}{2n}$. Notice that the composition of two such reflections is a rotation that belongs to C_n . There are $2n$ elements in D_n : namely n rotations $\rho, \rho^2, \dots, \rho^n = \iota$, and n reflections that can be expressed as $\rho\sigma, \rho^2\sigma, \dots, \rho^n\sigma = \sigma$, where σ is any one of the reflections.

Here are the cases with small $n = 1$ and 2 :

- $C_1 = \{\iota\}$ and $D_1 = \{\iota, \sigma_m\}$,
- $C_2 = \{\iota, \sigma_P\}$ and $D_2 = \{\iota, \sigma_P, \sigma_m, \sigma_l\}$, where m and l are perpendicular lines through point P .

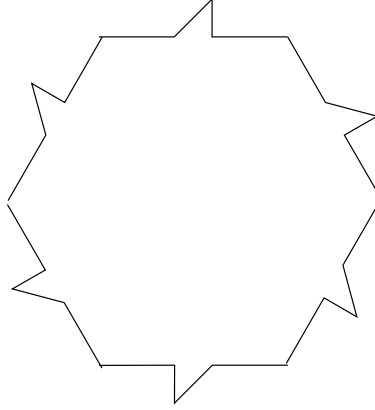
Example 7. The symmetry group of a polygon with n edges and vertices (called n -gon) can contain at most $2n$ elements. Indeed, any symmetry α must map vertices into vertices, and neighboring vertices into neighboring vertices. Fixed vertex A has at most n possible images. Adjacent vertex B then has at most two alternatives as it must be mapped into one of the two vertices next to $\alpha(A)$. After this, the symmetry is uniquely determined.

Let us show that the regular n -gon has exactly $2n$ symmetries, and the symmetry group is the dihedral group D_n . Let P be the center of the regular n -gon. It is clear that the rotation $\rho = \rho_P, \frac{360^\circ}{n}$ is a symmetry of the n -gon. If m is line through P and one of the vertices then also $\sigma = \sigma_m$ is a symmetry. As the symmetries form a group, all products of σ and ρ are symmetries. These include the n rotations

$\rho, \rho^2, \dots, \rho^n = \iota$ generated by ρ , and n distinct odd isometries $\rho\sigma, \rho^2\sigma, \dots, \rho^n\sigma = \sigma$. (These are distinct as $\rho^i\sigma = \rho^j\sigma \implies \rho^i = \rho^j$.) These are exactly the elements of D_n . There can be no other isometries as no n -gon can have more than $2n$ symmetries.

□

Example 8. Cyclic group C_n is the symmetry group of a polygon that is obtained from a regular n -gon by replacing each edge with a "directed edge", for example as follows:



□

Before proving that no other finite subgroups of $\mathcal{I}$ exists, let us first figure out multiplication rules of even isometries.

Theorem 2.21 1. *The product of two translations is a translation.*

2. *A rotation by angle Θ followed by a rotation by angle Φ is a rotation by angle $\Theta + \Phi$, unless $\Theta + \Phi$ is a multiple of 360° , in which case the product is a translation.*
3. *A translation followed by a non-trivial rotation by Θ is a rotation by Θ . Also, a non-trivial rotation by Θ followed by a translation is a rotation by Θ .*

Proof.

1. Trivial: it follows from the definition of translations that $\tau_A\tau_B = \tau_{A+B}$.
2. If the two rotations are about the same center P then the claim is trivial: $\rho_{P,\Theta}\rho_{P,\Phi} = \rho_{P,\Theta+\Phi}$. Assume then that the two rotations are about different points A and B . Let m be the line through points A and B . According to Theorem 2.11 there exist lines l and n through points A and B , respectively, such that

$$\rho_{A,\Theta} = \sigma_m\sigma_l \quad \text{and} \quad \rho_{B,\Phi} = \sigma_n\sigma_m,$$

so

$$\rho_{B,\Phi}\rho_{A,\Theta} = \sigma_n\sigma_m\sigma_m\sigma_l = \sigma_n\sigma_l.$$

Moreover, the directed angle from l to m is $\Theta/2$ and the directed angle from m to n is $\Phi/2$, so the directed angle from l to n is $(\Theta + \Phi)/2$. If this angle is a multiple of 180° then lines l and n are parallel, that is, if $\Theta + \Phi$ is a multiple of 360° then $\rho_{B,\Phi}\rho_{A,\Theta}$ is a translation. Otherwise lines l and n are not parallel, so $\rho_{B,\Phi}\rho_{A,\Theta}$ is a rotation by angle $\Theta + \Phi$.

3. Let τ be a translation and ρ a non-trivial rotation by angle Θ . Then $\tau = \sigma_l \sigma_m$ for parallel lines l and m , and $\rho = \sigma_n \sigma_k$ where the angle from line k to line n is $\Theta/2$. By Theorem 2.11 We can choose k to be parallel to l and m . Then $\rho\tau = \sigma_n \sigma_k \sigma_l \sigma_m$. Because k, l and m are parallel lines, by Corollary 2.10 the product $\sigma_k \sigma_l \sigma_m$ is a reflection σ_p where p is also parallel to k, l and m . The angle from line p to line n is $\Theta/2$, so $\rho\tau = \sigma_n \sigma_p$ is a rotation by angle Θ .

Analogously, we could have chosen k and n so that n is parallel to l and m , in which case $\sigma_l \sigma_m \sigma_n = \sigma_q$ for a line q in the same direction. Then

$$\tau\rho = \sigma_l \sigma_m \sigma_n \sigma_k = \sigma_q \sigma_k$$

is a rotation by angle Θ .

□

By iterating the theorem we easily get a rule for composing an arbitrary number of rotations:

$$\rho_{C_1, \Theta_1} \circ \rho_{C_2, \Theta_2} \circ \cdots \circ \rho_{C_n, \Theta_n}$$

is a rotation by angle $\Theta = \Theta_1 + \Theta_2 + \cdots + \Theta_n$, unless Θ is a multiple of 360° , in which case the product is a translation.

Corollary 2.22 *If a subgroup of $\mathcal{I}$ contains two non-trivial rotations about different centers then it also contains a non-trivial translation*

Proof. Let $\rho_{A, \Theta}$ and $\rho_{B, \Phi}$ be two non-trivial rotations and $A \neq B$. According to our theorem

$$\rho_{B, \Phi}^{-1} \rho_{A, \Theta}^{-1} \rho_{B, \Phi} \rho_{A, \Theta} = \rho_{B, -\Phi} \rho_{A, -\Theta} \rho_{B, \Phi} \rho_{A, \Theta}$$

is a translation. If it were the trivial translation ι then

$$\rho_{A, \Theta} \rho_{B, \Phi} = \rho_{B, \Phi} \rho_{A, \Theta}$$

but this is not possible as it was proved in a homework problem that non-trivial rotations about different centers do not commute.

□

Now we are ready to prove the result mentioned in the beginning of this section:

Theorem 2.23 (Leonardo da Vinci's Theorem) *A finite subgroup of $\mathcal{I}$ is either a cyclic group C_n or a dihedral group D_n .*

Proof. Let H be a finite subgroup of $\mathcal{I}$. Every non-trivial translation generates an infinite subgroup, so H cannot contain non-trivial translations. If γ is a glide reflection with glide vector A then γ^2 is a translation by vector $2A$, so H cannot contain any glide reflections except plain reflections. So only rotations and reflections are possible.

By the previous lemma, all rotations in H must have the same center P . Let $\rho = \rho_{P, \Theta}$ be the rotation having the smallest positive angle Θ among all rotations in H . It exists as H is finite. Let $\rho_{P, \Phi} \in H$. For every real number Φ there exists an integer k such that $0 \leq \Phi - k\Theta < \Theta$. Because the rotation by $\Phi - k\Theta$ is in H , and because Θ is the smallest positive angle, we must have $\Phi - k\Theta = 0$. This means that $\rho_{P, \Phi} = \rho^k$. We have proved that ρ generates the rotations of H . This means that the set of even isometries in H is $\{\rho, \rho^2, \dots, \rho^n\} = C_n$ for some n .

If there are no reflections in H then $H = C_n$. Assume then that there is at least one reflection σ in H . Then there are at least n distinct odd isometries $\sigma\rho, \sigma\rho^2, \dots, \sigma\rho^n$ in H . On the other hand, if $\alpha \in H$ is odd then $\sigma\alpha$ is even, so $\sigma\alpha = \rho^k$ for some $k = 1, 2, \dots, n$. This means that $\alpha = \sigma\rho^k$, and we have proved that $H = D_n$. □

Corollary 2.24 *The symmetry group of every polygon is a cyclic group or a dihedral group.*

Proof. In the example at the beginning of the section we concluded that the symmetry group of an n -gon contains at most $2n$ elements, so it is finite. □

2.8 Conjugacy

Two elements x and y of a group G are called conjugate if there exists an element $\alpha \in G$ such that $x = \alpha y \alpha^{-1}$. It is easy to see that conjugacy is an equivalence relation. Its equivalence classes are called the conjugacy classes of the group.

It turns out that in the group $\mathcal{I}$ conjugate isometries are of the same type (both translations, both rotations, both reflections or both glide reflections):

Theorem 2.25 *Let $\alpha \in \mathcal{I}$ be an arbitrary isometry.*

1. *Let $\sigma = \sigma_m$ be the reflection in line m . Then $\alpha\sigma\alpha^{-1}$ is the reflection $\sigma_{\alpha(m)}$ in line $\alpha(m)$.*
2. *Let $\tau = \tau_{B-A}$ be the translation that moves point A to point $B = \tau(A)$. Then $\alpha\tau\alpha^{-1}$ is the translation $\tau_{\alpha(B)-\alpha(A)}$ that moves point $\alpha(A)$ to point $\alpha(B)$.*
3. *Let $\rho = \rho_{P,\Theta}$ be a rotation about point P . Then $\alpha\rho\alpha^{-1}$ is the rotation $\rho_{\alpha(P),\pm\Theta}$ about point $\alpha(P)$, where the angle is $+\Theta$ if α is even, and $-\Theta$ if α is odd.*
4. *Let $\gamma = \gamma_{m,B-A}$ be a glide reflection. Then $\alpha\gamma\alpha^{-1}$ is the glide reflection $\gamma_{\alpha(m),\alpha(B)-\alpha(A)}$.*

Proof.

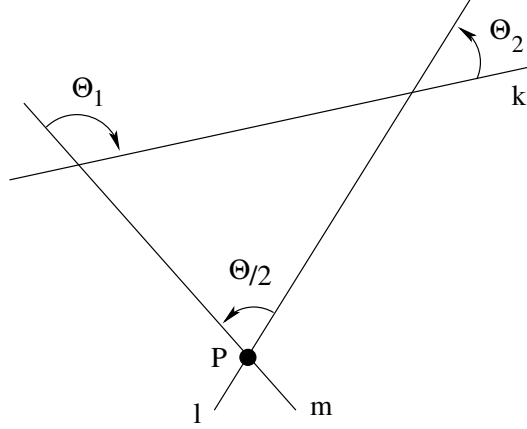
1. Isometry $\alpha\sigma_m\alpha^{-1}$ is an odd isometry that fixes every point $\alpha(P)$ of line $\alpha(m)$. The only odd isometry with this property is the reflection in line $\alpha(m)$.
2. Let $\tau = \tau_{B-A}$ be the translation that moves A to B . Then $\tau = \sigma_m\sigma_l$ for two parallel lines m and l . According to case 1 above, $\alpha\sigma_m\alpha^{-1} = \sigma_{\alpha(m)}$ and $\alpha\sigma_l\alpha^{-1} = \sigma_{\alpha(l)}$. We get

$$\alpha\tau\alpha^{-1} = \alpha\sigma_m\sigma_l\alpha^{-1} = \alpha\sigma_m\alpha^{-1}\alpha\sigma_l\alpha^{-1} = \sigma_{\alpha(m)}\sigma_{\alpha(l)}.$$

Isometries preserve parallelism of lines, so $\alpha(m)$ and $\alpha(l)$ are parallel lines, which means that $\alpha\tau\alpha^{-1}$ is a translation. It moves point $\alpha(A)$ into $\alpha\tau\alpha^{-1}\alpha(A) = \alpha(B)$ so it is the translation $\tau_{\alpha(B)-\alpha(A)}$.

3. Let $\rho = \rho_{P,\Theta}$ where $\Theta \neq 0$. (The case $\rho = \iota$ is trivial.) Clearly $\alpha\rho\alpha^{-1}$ is an even isometry with fixed point $\alpha(P)$, so $\alpha\rho\alpha^{-1}$ must be some rotation about point $\alpha(P)$, say $\alpha\rho\alpha^{-1} = \rho_{\alpha(P),\Phi}$. All we need to prove is that $\Phi = \pm\Theta$ where the sign depends on the parity of α .

Assume first that $\alpha = \sigma_k$ for some line k . Let m and l be lines through point P such that the directed angle from l to m is $\Theta/2$, so $\rho = \sigma_m\sigma_l$. We are free to choose lines m and l in such a way that neither is parallel to k . Let Θ_1 and Θ_2 be the directed angles from m to k and from k to l , respectively. Notice that $\Theta_1 + \Theta_2$ is then the directed angle from m to l , that is, $\Theta_1 + \Theta_2 = -\Theta/2$, at least modulo 180° .



We have

$$\rho_{\alpha(P),\Phi} = \alpha\rho\alpha^{-1} = \sigma_k\sigma_m\sigma_l\sigma_k.$$

This is the product of two rotations $\sigma_k\sigma_m$ and $\sigma_l\sigma_k$ of angles $2\Theta_1$ and $2\Theta_2$, respectively. According to Theorem 2.21 the product is a rotation by $2\Theta_1 + 2\Theta_2 = -\Theta$, that is, $\Phi = -\Theta$ as required.

Assume then a general α . We know that all isometries are products of (at most three) reflections, so $\alpha = \sigma_1\sigma_2\ldots\sigma_n$ for some reflections $\sigma_1, \sigma_2, \ldots, \sigma_n$. Number n is even iff isometry α is even. As

$$\alpha\rho\alpha^{-1} = \sigma_1\sigma_2\ldots\sigma_n\rho\sigma_n\sigma_{n-1}\ldots\sigma_1$$

we can apply the single reflection case n times. In each application the sign of the rotation angle changes, so in the end we have that $\alpha\rho\alpha^{-1}$ is a rotation by the angle $(-1)^n\Theta$.

4. Let $\gamma = \gamma_{m,B-A}$, where $A \neq B$. (If $A = B$ then γ is a reflection, and that was already taken care of.) Then $\alpha\gamma\alpha^{-1}$ is an odd isometry, so it is a glide reflection, say γ' . Because $\gamma'(\alpha(m)) = \alpha(m)$, line $\alpha(m)$ must be the axis of γ' . To find the glide vector of γ' we can make the calculation

$$\gamma'\gamma' = \alpha\gamma\alpha^{-1}\alpha\gamma\alpha^{-1} = \alpha\gamma^2\alpha^{-1} = \alpha\tau_{B-A}\tau_{B-A}\alpha^{-1} = \alpha\tau_{B-A}\alpha^{-1}\alpha\tau_{B-A}\alpha^{-1} = \tau_{\alpha(B)-\alpha(A)}\tau_{\alpha(B)-\alpha(A)},$$

which shows that $\alpha(B) - \alpha(A)$ is the glide vector of γ' .

□

Let $s \subseteq \mathbb{R}^2$. The following terminologies are widely used: If σ_m is a symmetry of s then m is called a line of symmetry for s . If σ_P is a symmetry of s then P is a point of symmetry for s . If $\rho_{C,\Theta}$ is a symmetry of s then C is a center of symmetry and, more precisely, if $\Theta = \frac{360^\circ}{n}$ then C is a center of n -fold symmetry.

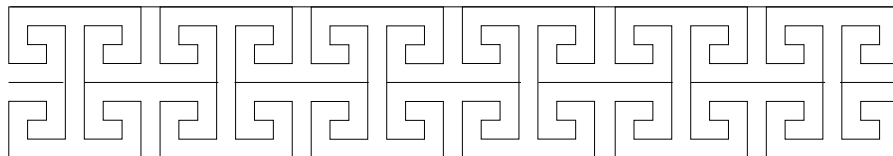
In analyzing symmetries we frequently apply the statements of the conjugacy theorem above in the following forms. Let α be an arbitrary symmetry of set s . Then in s :

- If m is a line of symmetry then also $\alpha(m)$ is a line of symmetry.
- If P is a point of symmetry then also $\alpha(P)$ is a point of symmetry.
- If C is a center of (n -fold) symmetry then also $\alpha(C)$ is a center of (n -fold) symmetry.

2.9 Frieze groups

Let us denote the set of translations by $\mathcal{T}$. It is easily seen to be a subgroup of $\mathcal{I}$. The intersection of two subgroups is also a subgroup, so for every subgroup G of $\mathcal{I}$, the set $G \cap \mathcal{T}$ that contains the translations of G is a subgroup of G , called the translation group of G .

We say that $G \subseteq \mathcal{I}$ is a frieze group if its translation group is cyclic and non-trivial, that is, if the translations are generated by a single translation $\tau \neq \iota$. The name comes from the fact that frieze groups are the symmetry groups of repetitive friezes (=ornamented bands on buildings) such as, for example



(where the pattern is repeated indefinitely in both directions). Notice that there must exist the shortest translation that keeps the frieze invariant — otherwise its symmetry group is not a frieze group. For example, the symmetry group of a horizontal line is not a frieze group as it contains all horizontal translations. It turns out that there are only seven different frieze groups (when we ignore the position, orientation and the size of the frieze) and each is the symmetry group of some $s \subseteq \mathbb{R}^2$.

In this section we make the following convention: The direction of the translations in the frieze group is called the horizontal direction, and the perpendicular direction is then the vertical direction. We start with the following key observation:

Lemma 2.26 *Let G be a subgroup of $\mathcal{I}$ such that all translations in G are horizontal, and assume that there is at least one non-trivial translation. (This includes all frieze groups, but also groups without a shortest translation.) Then there exists a horizontal line m such that all elements of G are products of reflections in vertical lines, possibly followed by the reflection σ_m in line m . These products are:*

- *horizontal translations,*
- *reflections in vertical lines,*
- *reflection σ_m in line m ,*
- *halfturns about points of line m , and*
- *glide reflections with axis m .*

Proof. Let $\tau \in G$ be a fixed non-trivial translation.

First, let us prove that all non-trivial rotations in G are halfturns. Let $\rho = \rho_{P,\Theta} \in G$ be arbitrary. Let $A = \tau(P)$, so $A \neq P$, and let $B = \rho(A)$. According to Theorem 2.25, $\rho\tau\rho^{-1}$ is the translation that moves point $\rho(P) = P$ to point $\rho(A) = B$. Translations τ and $\rho\tau\rho^{-1}$ are horizontal, so points A , P and B must be on the same line. This is possible only if ρ is the trivial rotation or the halfturn about P .

Next, let us prove that all reflections in G are in vertical and horizontal lines. Let $\sigma_l \in G$ be arbitrary, P a point of line l , $A = \tau(P)$, and $B = \sigma_l(A)$. According to Theorem 2.25, $\sigma_l\tau\sigma_l^{-1}$ is the translation that moves point $\sigma_l(P) = P$ to point $\sigma_l(A) = B$. Again, translations τ and $\sigma_l\tau\sigma_l^{-1}$ are horizontal, so points A , P and B must be on the same horizontal line. Either $A = B$, in which case A is on line l so l is horizontal, or $A \neq B$, in which case l is the perpendicular bisector of AB so l is vertical.

Finally, let us show that glide reflections of G are horizontal. Indeed, if $\gamma \in G$ is a glide reflection with a non-zero glide A , then γ^2 is the translation with the translation vector $2A$. Vector $2A$ is horizontal, so also the glide A is horizontal.

In the three paragraphs above we have shown that every element of G is a product of reflections in vertical and horizontal lines. As products of reflections in perpendicular directions commute (Corollary 2.12), every element of G is a product

$$\sigma_1 \sigma_2 \dots \sigma_v \sigma^{(1)} \sigma^{(2)} \dots \sigma^{(h)}$$

where each σ_i is a reflection in a vertical line, and each $\sigma^{(j)}$ is a reflection in a horizontal line. Moreover, as the product of three reflections in parallel lines is a reflection in a parallel line, we can reduce the number of reflections so that $v, h \leq 2$.

Next we prove that, in fact, $h \leq 1$. Assume the contrary: some

$$\alpha = \sigma_1 \sigma_2 \dots \sigma_v \sigma^{(1)} \sigma^{(2)} \in G$$

where the reflections $\sigma^{(1)}$ and $\sigma^{(2)}$ are in two different horizontal lines. If $v = 1$ then α is a glide reflection with a non-zero vertical glide, and if $v = 0$ or $v = 2$ then α is a translation in a direction that is not horizontal. These isometries do not exist in G , so we must have $h \leq 1$.

Moreover, the possible reflection $\sigma^{(1)}$ in a horizontal line must be in the *same* horizontal line m for all isometries of G . Namely, if G would contain two isometries $\alpha = \alpha' \sigma^{(1)}$ and $\beta = \beta' \sigma^{(2)}$ where α' and β' are products of reflections in vertical lines and $\sigma^{(1)}$ and $\sigma^{(2)}$ are reflections in two different horizontal lines then the product

$$\alpha\beta = \alpha' \sigma^{(1)} \beta' \sigma^{(2)} = \alpha' \beta' \sigma^{(1)} \sigma^{(2)}$$

would contradict the previous paragraph.

So we conclude that every element of G is a product of 0,1 or 2 reflections in vertical lines, or a product of 0,1 or 2 reflections in vertical lines followed by σ_m , the reflection in the horizontal axis m of the group. This leaves the following non-trivial possibilities:

- σ_m : the reflection in the axis m ,
- $\sigma_1 \sigma_m$: a halfturn about a point of line m ,
- $\sigma_1 \sigma_2 \sigma_m$: a glide reflection with axis m ,
- σ_1 : a reflection in a vertical line, and
- $\sigma_1 \sigma_2$: a horizontal translation

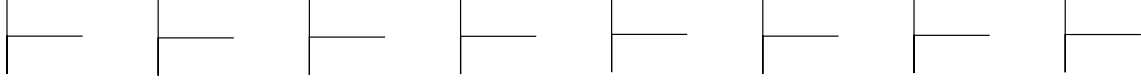
□

Now we are ready to classify all frieze groups. Let G be a frieze group whose translations are generated by the shortest translation τ_A , and let m be the horizontal line from the previous lemma, called the axis of the frieze group. Let $2d$ be the length of vector A , so that τ_A is a product of two reflections in vertical lines at distance d . The translations in G are then exactly the products of two reflections in any two vertical lines whose distance is a multiple of d .

1) Assume first that $\sigma_m \in G$. Let l and k be arbitrary vertical lines. Then

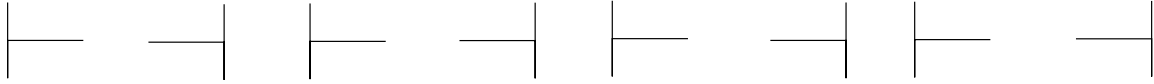
$$\begin{aligned} \sigma_l \sigma_m \in G &\iff \sigma_l \in G, \text{ and} \\ \sigma_l \sigma_k \sigma_m \in G &\iff \sigma_l \sigma_k \in G, \end{aligned}$$

so glide reflections of G are uniquely determined by the translations, and the reflections in vertical lines are uniquely determined by the halfturns. If there are no halfturns in G then G is generated by τ_A and σ_m , and it is the symmetry group of the infinite strip



Let us call it the group F_{1001} . The four indices 1001 are interpreted as follows: the group contains the reflection σ_m in the axis m , does not contain any reflection σ_l in vertical lines, does not contain any halfturns, and contains glide reflections. This information uniquely specifies the group (for given A and m).

Assume then that there is some halfturn σ_P in G . We know that point P is on line m . Because $\sigma_P \sigma_Q = \tau_{2(P-Q)}$ is a translation, the other halfturns are now uniquely determined: they are at points of line m whose distance from P is a multiple of d . These then also uniquely determine the reflections σ_l in vertical lines. Group $G = F_{1111}$ is the symmetry group of



Groups F_{1001} and F_{1111} are the only groups containing the reflection σ_m .

2) Consider then groups that do not contain σ_m . One alternative is that there are no isometries except the translations: We have the symmetry group F_{0000} of the strip



Let us assume then that there are other symmetries. The product of a reflection in a vertical line and a halfturn is a glide reflection, the product of a glide reflection and a halfturn is a reflection in a vertical line, and the product of a glide reflection and a reflection in a vertical line is a halfturn. Conclusion: G either contains all three types of isometries, or at most one of the types. There are four alternatives, resulting in groups F_{0100} , F_{0010} , F_{0001} and F_{0111} , as discussed below.

If G contains a halfturn $\sigma_P = \sigma_l \sigma_m$ then the other halfturns are uniquely determined: they are the products of σ_P and the translations in G . The distances between the centers of the halfturns are then exactly the multiples of d . This means that group F_{0010} is uniquely determined, and it is the symmetry group of the following strip:



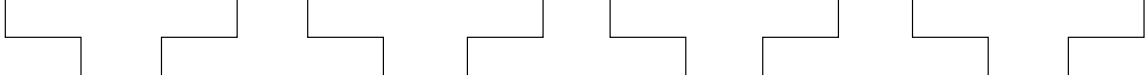
Analogously, if G contains a reflection σ_l in a vertical line l then the other reflections are determined as they must be the products of σ_l and the translations in G . The lines of the reflections are at distances that are multiples of d . So we have the group F_{0100} which is the symmetry group of



Consider then a glide reflection $\gamma = \sigma_l \sigma_k \sigma_m \in G$ with axis m . Let $2g$ be the length of its glide vector, that is, g is the distance between lines l and k . Then g must be a multiple of $d/2$ as γ^2 is a translation of length $4g$. On the other hand, g cannot be a multiple of d because then there would exist a translation in G that would cancel the glide, leaving σ_m , and we assumed that σ_m is not in G . We conclude that g must be an odd multiple of $d/2$, or equivalently, the length $2g$ of the glide is an odd multiple of d . All such glide reflections are obtained from γ by multiplying it with translations, so we have completely characterized the glide reflections. Group F_{0001} is the symmetry group of



The last open possibility is that G contains halfturns, reflections in vertical lines and glide reflections. As discussed above, the glide reflections are uniquely determined (the glides are by odd multiples of d), and after we fix one center P of a halfturn, also the halfturns are uniquely determined. This also fixes the reflections as they are the products of the glide reflections and the halfturns. The lines of the reflections bisect the consecutive points of reflections. We have the group F_{0111} , which is the symmetry group of the following strip:



We have fully classified the frieze groups, and we found seven different types. In each case, a "frieze" with the given symmetries was given, to prove that the seven types of frieze groups are the symmetry groups of some sets $s \subseteq \mathbb{R}^2$. Notice that each of the seven groups has infinitely many "geometric realizations", as the axis m can be any line, the shortest translation τ can be any non-trivial translation parallel to m , and in those groups that involve halfturns or reflections in vertical line, one center P of a halfturn or one line l of a reflection can be selected. But modulo these parameters, the groups are unique. It is clear that all realizations of each group are isomorphic, and even more than that, isomorphic by isomorphisms that preserve the type of isometry (translations correspond to translations, reflections to reflections, rotations to rotations, ...).

We have proved the following theorem:

Theorem 2.27 *Let G be a frieze group whose translations are generated by τ . Then there exists a line m parallel to τ , and if G contains a halfturn there exists a point $P \in m$, otherwise a line l perpendicular to m , such that G is one of the following seven groups:*

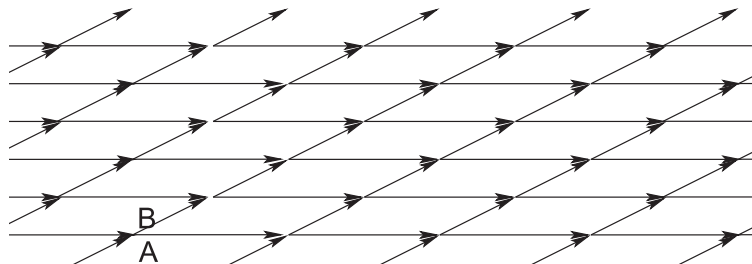
$$\begin{array}{lll} F_{0000} = \langle \tau \rangle & F_{1001} = \langle \tau, \sigma_m \rangle & F_{1111} = \langle \tau, \sigma_m, \sigma_P \rangle \\ F_{0100} = \langle \tau, \sigma_l \rangle & F_{0010} = \langle \tau, \sigma_P \rangle & \\ F_{0001} = \langle \gamma \rangle & F_{0111} = \langle \gamma, \sigma_P \rangle & \end{array}$$

where γ is the glide reflection with axis m such that $\gamma^2 = \tau$.

□

2.10 Wallpaper groups

A wallpaper group G is a subgroup of $\mathcal{I}$ whose translations are generated by two non-parallel translations τ_1 and τ_2 . Translations commute with each other, so the translations of G are exactly the isometries $\tau_1^i \tau_2^j$ for all integers i and j . If A and B are the vectors of translations τ_1 and τ_2 then the vectors of translations $\tau_1^i \tau_2^j$ are $iA + jB$, which form a lattice



Let us first show that there exists a shortest translation in G .

Lemma 2.28 *Wallpaper group G has a shortest non-trivial translation. More generally, any non-empty subset s of translations of G contains a shortest non-trivial translation.*

Proof. Let A and B be the translation vectors of the generating translations τ_1 and τ_2 . Let

$$B = rA + B' \quad \text{and} \quad A = qB + A'$$

be the decompositions of vectors A and B into a sum of orthogonal vectors, where $r, q \in \mathbb{R}$ and $B' \perp A$ and $A' \perp B$. As A and B are not parallel, vectors A' and B' are non-zero. Let $a > 0$ and $b > 0$ be the lengths of vectors A' and B' , respectively.

Consider an arbitrary translation vector $A_{ij} = iA + jB$ in G . Using the orthogonal decompositions above we have

$$A_{ij} = (i + jr)A + jB' \quad \text{and} \quad A_{ij} = (j + iq)B + iA'.$$

These are sums of two orthogonal vectors, so the length of A_{ij} is at least $|j|b$, the length of jB' , and at least $|i|a$, the length of iA' . Let c be the length of some vector X in the set s of translations we consider. Then any vector A_{ij} with $|j| > c/b$ or $|i| > c/a$ is longer than vector X . Therefore there are only a finite number of vectors that can potentially be shorter than X . The shortest among them is the shortest translation vector in set s . $\square$

Rosette groups, frieze groups and wallpaper groups are exactly the discrete symmetry groups: We call a subgroup G of $\mathcal{I}$ discrete if it does not contain arbitrarily short translations and does not contain arbitrarily small rotations. More precisely, G is discrete if there exists $\varepsilon > 0$ such that

$$\begin{aligned} 0 < |A| < \varepsilon &\implies \tau_A \notin G, \text{ and} \\ 0 < \Theta < \varepsilon &\implies \rho_{C,\Theta} \notin G. \end{aligned}$$

($|A|$ is the length of the translation vector A .)

Theorem 2.29 *Discrete subgroups of $\mathcal{I}$ are exactly the rosette groups, frieze groups and wallpaper groups.*

Proof. ($\Leftarrow$) Rosette groups are finite and hence discrete. In frieze groups, the translation that generates all translations is the shortest one, and halfturns are the only possible rotations, so frieze groups are discrete. Let G be a wallpaper group. By Lemma 2.28, it contains a shortest translation τ . For every rotation $\rho \in G$, the isometry $\tau' = \rho\tau\rho^{-1}$ is the translation that maps the center C of ρ to $\rho\tau(C)$. Consequently, translation $\tau'\tau^{-1}$ takes point $\tau(C)$ into $\rho\tau(C)$. This translation is arbitrarily short for arbitrarily small rotation angles, so G cannot contain arbitrarily small rotations. Hence G is discrete.

($\Rightarrow$) Let G be a discrete subgroup of $\mathcal{I}$.

(1) If G contains no non-trivial translations then it does not contain any glide reflections with non-zero glide vector. There are only rotations and reflections in G . Rotations can only have a finite number of different rotation angles as otherwise there would be arbitrarily small rotations in G . Two rotations by the same angles but with different centers generate a translation, so the rotations of G have the same center C . Reflection lines must contain C , and there are only a finite number of possible angles between the lines of reflections. We conclude that the group is finite, and hence it is a rosette group.

(2) Suppose then that G contains a non-trivial translation τ_A . Due to discreteness there can be only a finite number of different translations by vectors shorter than A , so a shortest non-trivial translation exists. We may assume τ_A is a shortest translation.

(2a) If all translations in G are generated by τ_A then G is a frieze group.

(2b) Suppose then that there exists a translation τ_B in G that is not generated by τ_A . Again, by discreteness, a shortest such translation exists, so we may assume that B has minimum length. To complete the proof of Theorem 2.29, we use the following lemma that states that τ_A and τ_B generate all translations in G , implying that G is a wallpaper group.

Lemma 2.30 *Let G be a discrete subgroup of $\mathcal{I}$, let τ_A be a shortest non-zero translation in G , and let $\tau_B \in G$ be a shortest translation not generated by τ_A . Then τ_A and τ_B generate all translations of G .*

Proof. It is clear that vectors A and B are not in parallel directions (otherwise A would not be the shortest translation vector), so every vector of $\mathbb{R}^2$ is a linear combination of A and B . Assume that group G contains a translation τ_C such that $\tau_C \notin \langle \tau_A, \tau_B \rangle$. Let $C = xA + yB$ be the representation of C as a linear combination of vectors A and B , where $x, y \in \mathbb{R}$. By subtracting integer multiples of vectors A and B from vector C , we can reduce x and y so that $-\frac{1}{2} \leq x, y \leq \frac{1}{2}$. But then, using the triangular inequality, we obtain

$$|C| = |xA + yB| \leq |x||A| + |y||B| \leq (|A| + |B|)/2 \leq |B|.$$

The first inequality can be an equality only if $x = 0$ or $y = 0$, but in these cases the second inequality is proper. So in each case: $|C| < |B|$, which contradicts the minimality of vector B . $\square$

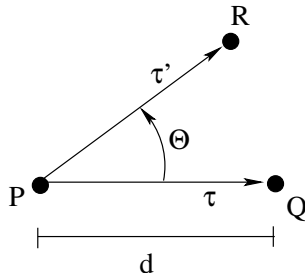
Let us start analyzing the possibilities for the wallpaper groups. It turns out that there are 17 different types of groups. Deriving them is a lengthy case analysis. The rest of this chapter provides a complete derivation. (See also the slide presentation from the course web page.)

Our first observation is an important restriction on possible rotations in wallpaper groups:

Theorem 2.31 (Crystallographic restriction) *A wallpaper group G can only contain rotations by multiples of 60° and 90° . Hence all centers of rotations are centers of n -fold rotations for $n = 2, 3, 4$ or 6 . Moreover, a 4-fold rotation cannot co-exist with 3- or 6-fold rotations.*

Proof. Let $\tau = \tau_A$ be the shortest translation in G , and let d be the length of its translation vector A .

Let $\rho = \rho_{P,\Theta} \in G$ be a non-trivial rotation, and let $Q = \tau(P)$ and $R = \rho(Q)$. Then G contains also the translation $\tau' = \rho\tau\rho^{-1}$ that moves point $\rho(P) = P$ to point $\rho(Q) = R$. Translation $\tau'\tau^{-1}$ then moves point Q to point R .



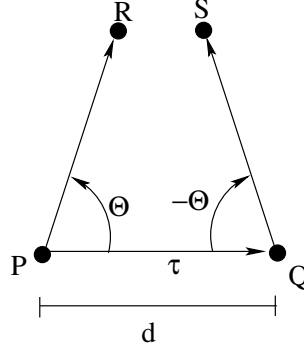
If $0^\circ < \Theta < 60^\circ$ then the distance between points Q and R is less than d , which contradicts the fact that τ is the shortest translation in G . We conclude that every non-trivial rotation is by an angle that is at least 60° . This also means that G can contain at most 6 different rotations about point P , because if we would have rotations by angles $\Theta_1, \Theta_2, \dots, \Theta_7$ where

$$0^\circ \leq \Theta_1 < \Theta_2 < \dots < \Theta_7 < 360^\circ$$

then necessarily $0^\circ < \Theta_{i+1} - \Theta_i < 60^\circ$ for some $i = 1, 2, \dots, 6$, a contradiction.

Let Θ be the smallest positive rotation angle about point P , and let Φ be any other rotation angle about P . There exists an integer k such that $0 \leq \Phi - k\Theta < \Theta$. This implies that $\Phi = k\Theta$. Therefore the rotations about point P are generated by $\rho_{P,\Theta}$, and $\Theta = \frac{360^\circ}{n}$ for some $n \leq 6$.

We still have to show that the case $n = 5$ of five-fold rotations is not possible. The rotation angle of a five-fold rotation is $\Theta = 72^\circ$. Consider points P, Q and R as in the beginning of the proof. Point Q is the center of rotation $\tau\rho\tau^{-1}$ by the same angle Θ , and therefore G contains the rotation of $-\Theta$ about Q . Let $S = \rho'(P)$.



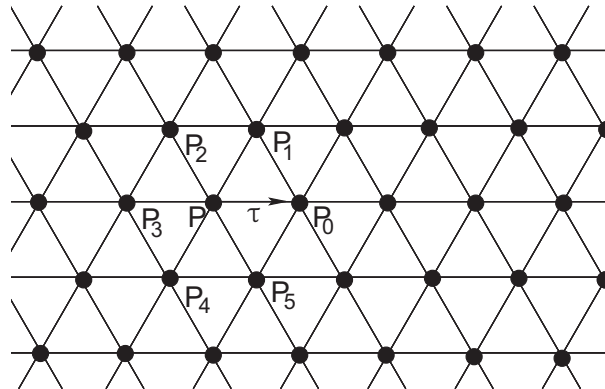
It is easily seen that the distance between points R and S is positive but less than d for angles in the interval $60^\circ < \Theta < 90^\circ$. In particular, this includes the case $\Theta = 72^\circ$ of five-fold rotations. Since G must contain the translation that moves R to S , this contradicts the minimality of distance d .

Finally we easily observe that if G contains a rotation ρ of 90° then it cannot contain any rotation ρ' of 60° or 120° because $\rho^{-1}\rho'$ would be a rotation whose angle is $\pm 30^\circ$.

□

Let us start analyzing different wallpaper groups case-by-case depending on the largest order of rotation that G contains.

1) Assume that G contains a 6-fold rotation $\rho = \rho_{P,60^\circ}$. Let τ be the shortest translation in G , let d be its length, and let $P_0 = \tau(P)$. Rotating point P_0 about point P defines points $P_i = \rho^i(P_0)$ for $i = 1, 2, \dots, 5$ such that all translations τ_{P_i-P} are in G . Then each P_i is a center of a 6-fold rotation in G . These isometries are all generated by ρ and τ through conjugacies. We can repeat the reasoning on all P_i , and then again on the six centers of rotation around them and so on. We conclude that G contains 6-fold rotations about centers that are the vertices of a lattice of equilateral triangles, and G contains all translations between vertices of the lattice. Let us denote by s_6 the set of the lattice points, indicated by black circles in the following figure:

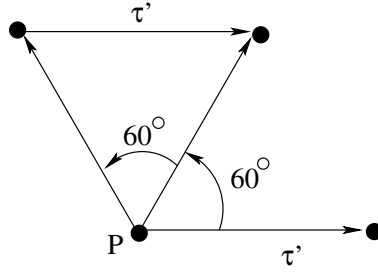


Let us show that the even isometries in G are exactly the even symmetries of s_6 . First, there can be no translation that moves a lattice point into a non-lattice point: The distance from every point of the plane to the closest lattice point is less than d , so if τ' is a translation that moves lattice point P into a non-lattice point $Q = \tau'(P)$ then the translation that moves Q to its closest lattice point is in G and it is shorter than τ , which contradicts the minimality of τ . So the translations of G are exactly that translations that keep s_6 invariant.

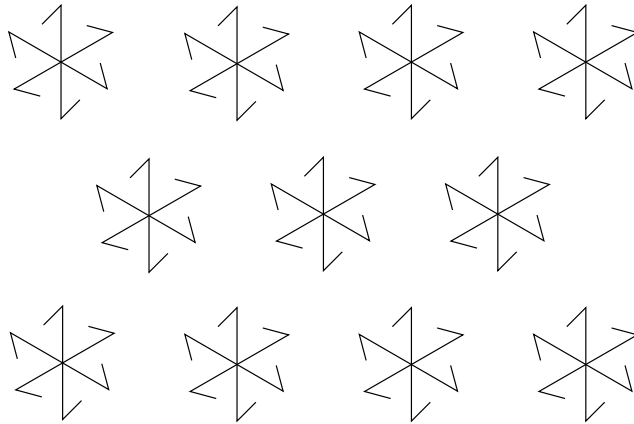
Consider then an arbitrary rotation $\rho' \in G$. The crystallographic restriction states that ρ' is a 2-, 3- or 6-fold rotation. This means that $\rho'\rho'^i$ is a translation for some integer i . Since translations in G are symmetries of s_6 , and since ρ is a symmetry of s_6 we conclude that ρ' is also a symmetry of s_6 .

Conversely, if ρ' is any rotation in the symmetry group of s_6 then it must be a 2-, 3- or 6-fold rotation (as the symmetry group of s_6 is a wallpaper group that contains 6-fold rotations) so $\rho'\rho'^i$ is a translation for some integer i . As ρ is a symmetry of s_6 this translation is also a symmetry of s_6 . All such translations are in G , so $\rho' \in G$ as well.

We have proved that the even elements of G are exactly the even symmetries of s_6 . If there are no odd isometries in G we have our first wallpaper group $W_6 = \langle \tau, \rho_{P,60^\circ} \rangle$ that consists of the even symmetries of s_6 . In addition to the translations and 6-fold rotation about lattice points this group also contains 3-fold rotations about the centers of the equilateral triangles, and 2-fold rotations about the midpoints between adjacent lattice points. Notice that the lattice points are the only centers of 6-fold rotations, because if ρ' is a 60° rotation then $\rho'\rho^{-1} = \tau'$ is a translation and, since $\tau'\rho\rho\tau'(P) = \rho\tau'(P)$, the lattice point $\rho\tau'(P)$ is the fixed point of $\rho' = \tau'\rho$.

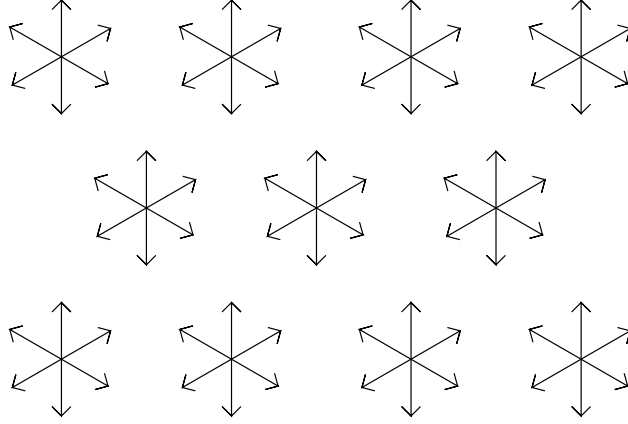


Group W_6 is the symmetry group of the following pattern where odd isometries are prevented by "directing" the lattice points counter-clockwise:

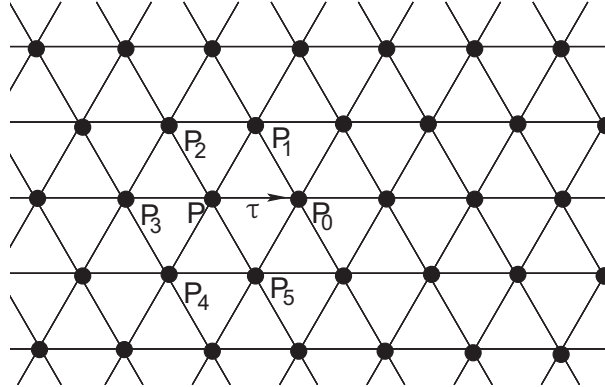


Assume then that G also contains some odd isometry α . This isometry has to take 6-fold rotation centers of G into 6-fold rotation centers of G , that is, α is a symmetry of s_6 . If β is any other odd symmetry of s_6 then $\alpha\beta \in G$ as $\alpha\beta$ is an even symmetry of s_6 , so also $\beta \in G$. Conclusion: G is the symmetry group of

s_6 . Note that s_6 has odd symmetries (e.g. a reflection σ in any line through two closest lattice points), so we have a new wallpaper group $W_6^1 = \langle \tau, \rho_{P,60^\circ}, \sigma \rangle$. Set s_6 is an example of a pattern whose symmetry group is W_6^1 . Here is another one:

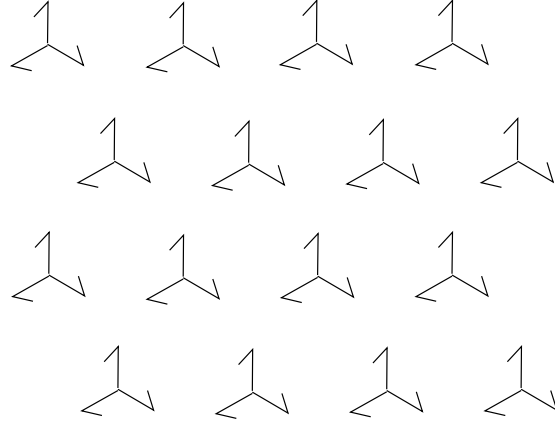


2) Assume that G contains a 3-fold rotation $\rho = \rho_{P,120^\circ}$ but no 6-fold rotations. We start in the same way as with the 6-fold rotations: Let τ be the shortest translation in G , let d be its length, and let $P_0, P_1, \dots, P_5$ be the points where P is taken by the translations $\tau, \rho^{-1}\tau^{-1}\rho, \rho\tau\rho^{-1}, \tau^{-1}, \rho^{-1}\tau\rho$ and $\rho\tau^{-1}\rho^{-1}$, respectively. Points $P_0, P_1, \dots, P_5$ are the vertices of the regular hexagon with center P , and they are all centers of 3-fold rotations in G . We can repeat the reasoning on each P_i instead of P , so we obtain again a lattice of equilateral triangles such that the vertices of the lattice are centers of 3-fold rotations, and the translations that move lattice points to lattice points are in G .



As before, let s_6 be the set of vertices of this lattice. Next we show that the even isometries of G are exactly those symmetries of s_6 that are translations or 3-fold rotations. First, exactly as in the case of W_6 , we see that no other translation is possible: a translation that moves a lattice point into a non-lattice point contradicts the minimality of translation τ . So the translations of G are exactly the translations that keep s_6 invariant. Consider then a rotation in G . We assumed that there are no 6-fold rotations, and therefore there can be no 2-fold rotations either (together with a 3-fold rotation any 2-fold rotation generates a 6-fold rotation). All other rotations would contradict the crystallographic restriction, so all rotations in G are 3-fold. Conversely, every 3-fold rotation ρ' that keeps s_6 invariant must be in G because $\rho'\rho^{-1}$ is a translation that keeps s_6 invariant, and all such translations are in G .

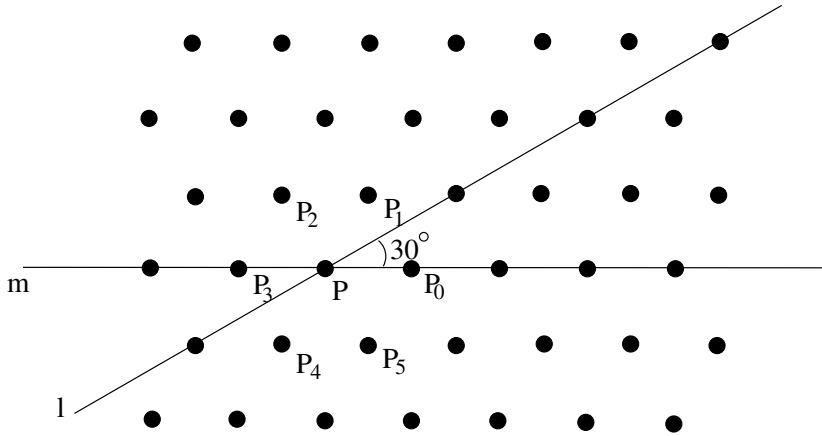
If there are no odd isometries in G we have our third wallpaper group $W_3 = \langle \tau, \rho_{P,120^\circ} \rangle$. In addition to the translations and 3-fold rotations about lattice points, group W_3 also contains 3-fold rotations about the centers of the equilateral triangles of the lattice. Group W_3 is the symmetry group of the following pattern:



Assume then that G also contains odd isometries. If G contains a glide reflection γ then it also contains a reflection because $\gamma\rho\gamma\rho^{-1}\gamma$ is a reflection for every glide reflection γ and 3-fold rotation ρ (homework). Every line p of reflection must contain a center of 3-fold rotation because also $\rho(p)$ is a line of reflection, lines p and $\rho(p)$ are not parallel so they intersect, and the product $\sigma_p\sigma_{\rho(p)}$ is a rotation about the point of intersection. In the beginning of case 3 the first center P of the 3-fold rotation ρ was chosen arbitrarily, so we may assume that P is on line p . Consequently P is a fixed point of a reflection in G .

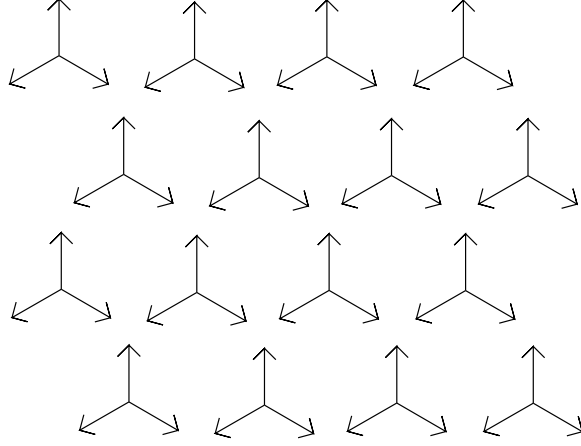
It follows then that every odd isometry in G is a symmetry of s_6 . Assume the contrary: there is an odd $\alpha \in G$ and a lattice point Q such that $\alpha(Q)$ is not a lattice point. Then $\alpha\tau_{Q-P}\sigma_p \in G$ is an even isometry that moves point P into the non-lattice point $\alpha(Q)$, and this contradicts the fact that all even isometries in G are symmetries of s_6 .

Let m be a line through two adjacent lattice points P and P_0 , and let l be the line through P such that the angle from m to l is 30° .

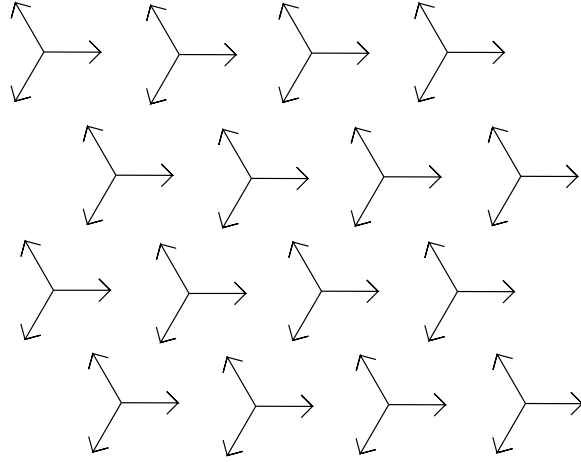


Both σ_m and σ_l are symmetries of s_6 , but since $\sigma_l\sigma_m$ is a rotation by 60° they cannot both be in group G . Let us prove that G must contain one of them. Assume the contrary: neither σ_m nor σ_l is in G , and let α be some odd isometry in G . Then $\alpha\sigma_m$ and $\sigma_l\alpha^{-1}$ are even symmetries of s_6 that do not belong to G , so they have to be rotations by an angle that is an odd multiple of 60° (=by $60, 180$ or -60 degrees). Their product $\sigma_l\alpha^{-1}\alpha\sigma_m = \sigma_l\sigma_m$ would then be a translation or a rotation by an even multiple of 60° , but we know that $\sigma_l\sigma_m$ is a rotation by 60° , a contradiction. We conclude that exactly one of the reflections σ_m and σ_l is in G .

Once we know one odd element of G , all other odd elements are uniquely determined by the even elements of G . We have two new wallpaper groups: $W_3^1 = \langle \tau, \rho_{P,120^\circ}, \sigma_l \rangle$, which is the symmetry group of



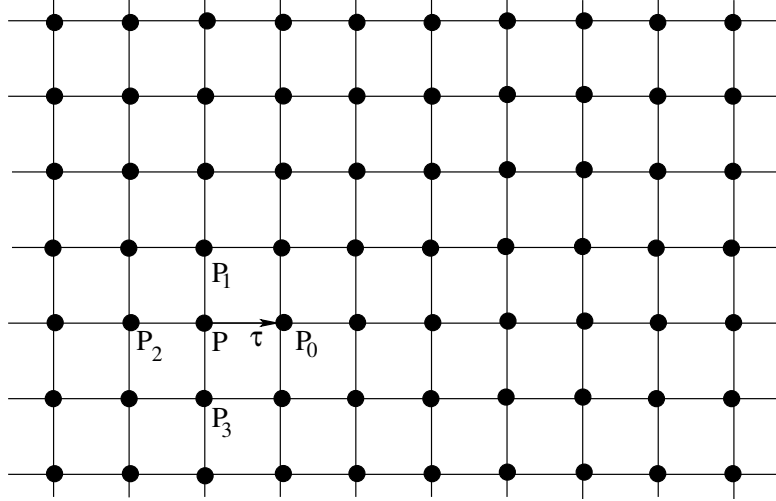
and $W_3^2 = \langle \tau, \rho_{P,120^\circ}, \sigma_m \rangle$, which is the symmetry group of



A difference between these groups is that W_3^1 contains a line of reflection through every center of 3-fold rotation, while in W_3^2 there are lines of symmetry only through some of the rotation centers, namely those that are the lattice points.

3) Let us assume now that G contains a 4-fold rotation $\rho_{P,90^\circ}$. Then it cannot contain 3- or 6-fold rotations. As in the previous cases: let τ be the shortest translation in G , let d be its length, and let P_i be the point where P is taken by the translation $\rho^i \tau \rho^{-i}$, for $i = 0, 1, 2$ and 3 . Points P_0, P_1, P_2 and P_3 are all centers of 4-fold rotations in G , so we can repeat the reasoning on each P_i . We obtain an infinite lattice of centers of 4-fold rotations, but this time the lattice is a square lattice instead of a triangular one. (See the next figure.) All translations between lattice points are in group G .

If G would contain any other translations, then it would contain a translation that moves a non-lattice point into the closest lattice point. This is not possible as the distance of every point of the plane from the lattice is less than d , the length of the shortest translation. We conclude that the translations in G are exactly the translations that keep the lattice invariant. Let us denote the points of the square lattice by s_4 .

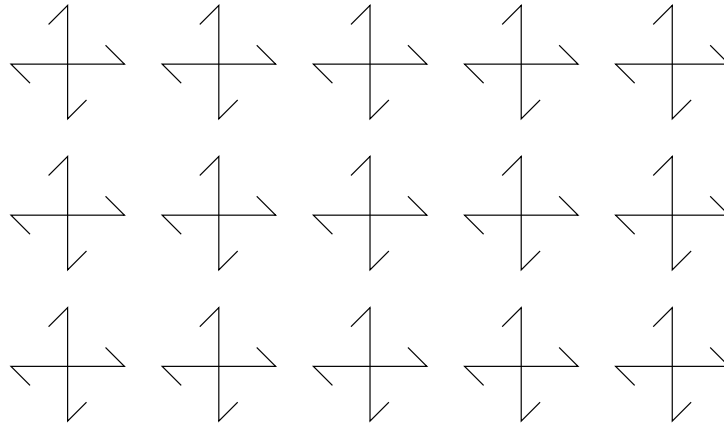


Analogously to the case of 60° rotations, we can prove that the even isometries in G are exactly the even symmetries of s_4 . We already know this for translations. Consider then an arbitrary rotation $\rho' \in G$. The crystallographic restriction states that ρ' is a 2- or 4-fold rotation. This means that $\rho'\rho^i$ is a translation for some integer i . Since translations in G are symmetries of s_4 , and since ρ is a symmetry of s_4 we conclude that ρ' is also a symmetry of s_4 .

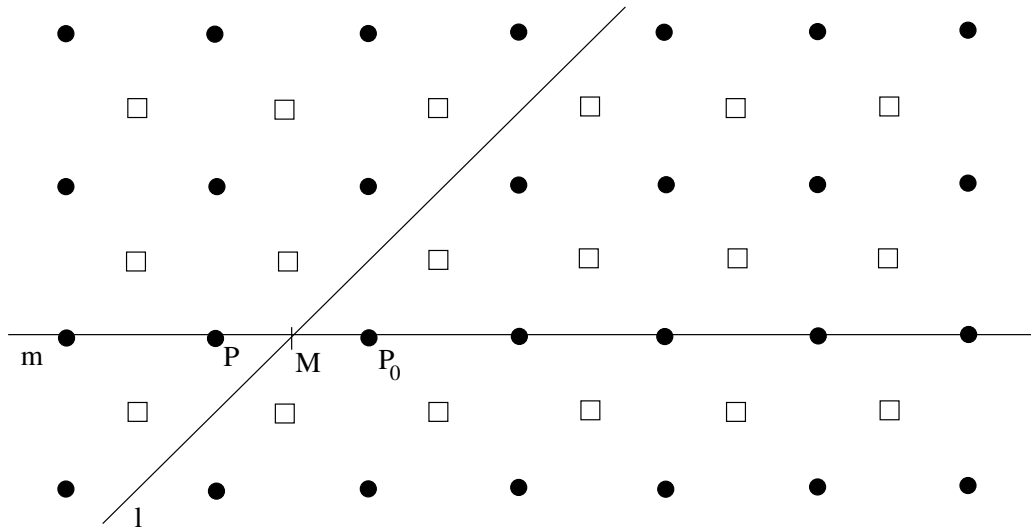
Conversely, if ρ' is any rotation in the symmetry group of s_4 then it must be a 2- or 4-fold rotation. This follows from the crystallographic restriction and the fact that the symmetry group of s_4 is a wallpaper group that contains 4-fold rotations. So $\rho'\rho^i$ is a translation for some integer i and, as ρ is a symmetry of s_4 , this translation is also a symmetry of s_4 . All such translations are in G , so $\rho' \in G$ as well.

If G contains no odd isometries then G is the group of even symmetries of s_4 . This is a new wallpaper group $W_4 = \langle \tau, \rho_{P,90^\circ} \rangle$. In addition to the translations and 4-fold rotations about lattice points this group also contains 4-fold rotations about the centers of the lattice squares, and 2-fold rotations about the midpoints between adjacent lattice points. Let us prove that no other rotations exist in G . Consider a center Q of a halfturn. Lattice point P is also a center of a halfturn. The product of the two halfturns is the translation by vector $2(Q - P)$. Translations are between lattice points, so Q must be a midpoint between lattice points. The only such points are the centers of the lattice squares (which are easily seen to be also centers of 4-fold rotations), and the midpoints between adjacent lattice points (which are easily seen not to be centers of 4-fold rotations). No other rotations are possible.

Group W_4 is the symmetry group of the following pattern:

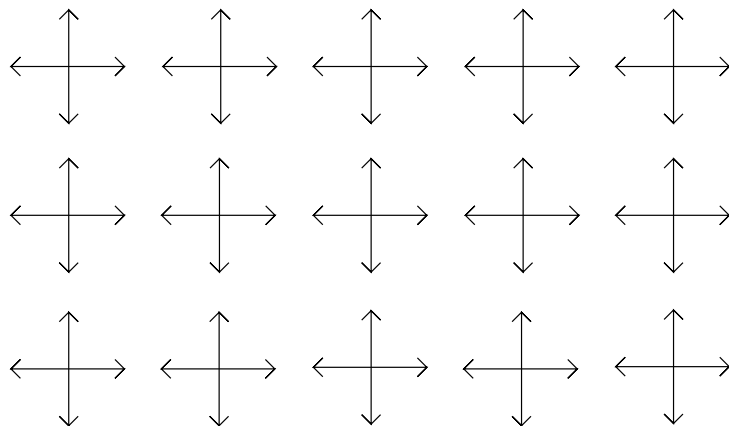


Assume then that G also contains odd isometries. Let m be a line through some adjacent lattice points, and let l be a line that intersects m at 45° in some midpoint M between adjacent lattice points:

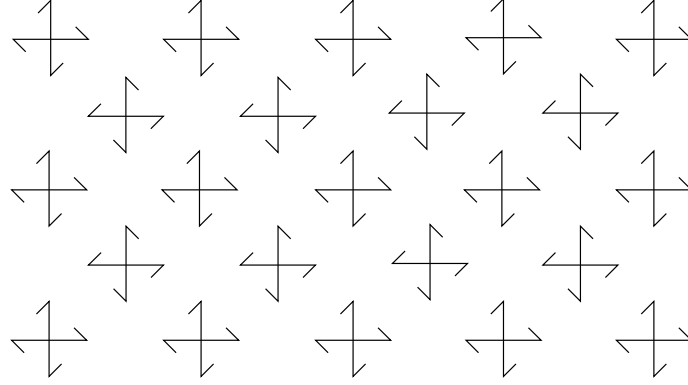


Reflection σ_m is a symmetry of s_4 whereas reflection σ_l is not. Instead, σ_l exchanges lattice points and the centers of the lattice squares. Group G cannot contain both σ_m and σ_l because then it would also contain a 4-fold rotation about point M . Let us prove that G must contain either σ_m or σ_l . If there is an odd isometry $\alpha \in G$ that takes some lattice point into a lattice point then every odd isometry of G must be a symmetry of s_4 . (Otherwise there would be an even element in G that is not a symmetry of s_4 .) As G contains all even symmetries of s_4 then all odd symmetries of s_4 are in G as well, and this includes σ_m . If, on the other hand, G contains an odd isometry α that takes all lattice points into non-lattice points then these non-lattice points must be the centers of the lattice squares, so $\sigma_l \alpha$ is an even symmetry of s_4 . Therefore $\sigma_l \alpha \in G$, and also $\sigma_l \in G$.

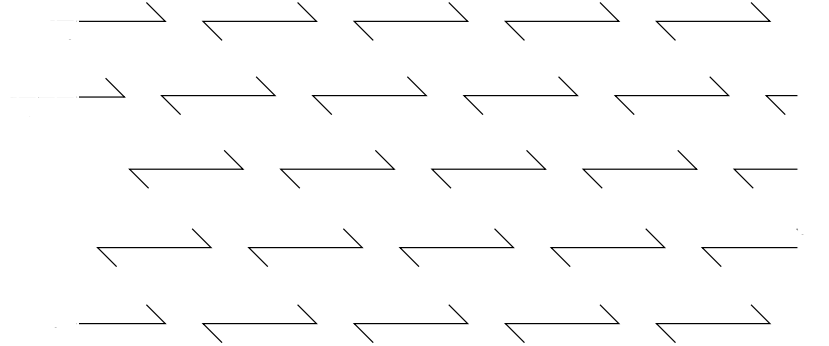
We have two new wallpaper groups $W_4^1 = \langle \tau, \rho_{P,90^\circ}, \sigma_m \rangle$, which is the symmetry group of



and $W_4^2 = \langle \tau, \rho_{P,90^\circ}, \sigma_l \rangle$, which is the symmetry group of



4) Assume that G contains halfturn σ_P , and that all non-trivial rotations in G are halfturns. Let τ_1 and τ_2 be two translations that generate all translations of G . Let the lattice points be the points $\tau_1^i \tau_2^j(P)$ for all integers i, j . They are all centers of halfturns. Also the products of σ_P and the translations $\tau_1^i \tau_2^j$ are halfturns about points that are midpoints between lattice points, that is, centers of the lattice parallelograms as well as the midpoints of their sides. No other halfturns are possible as otherwise we would get translations that are not invariants of the lattice. We conclude that we have found all even isometries in G . If G contains no odd isometries then we have the wallpaper group $W_2 = \langle \tau_1, \tau_2, \sigma_P \rangle$. It is the symmetry group of



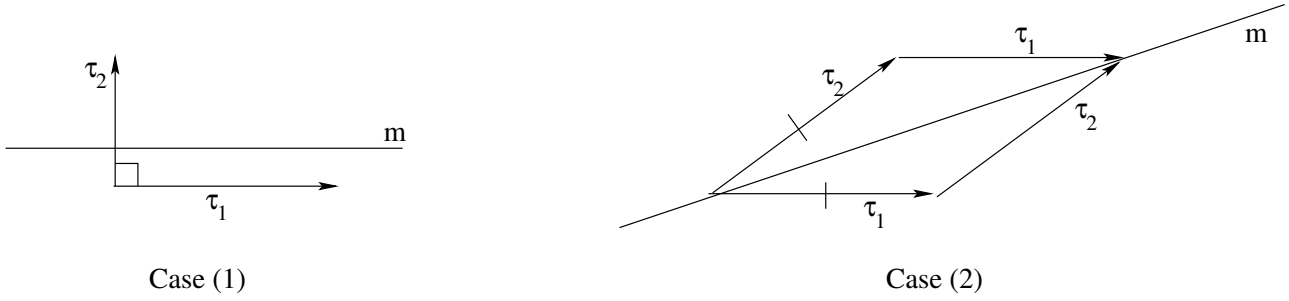
Assume then that G contains also some odd isometries. As in the previous cases, a single odd isometry $\alpha \in G$ uniquely determines all odd isometries because they are obtained by multiplying α with the even elements of G . The purpose of the following lemma is to limit the possible odd isometries that any wallpaper group can contain. It turns out that if G contains odd isometries then the translation lattice is rhombic or rectangular:

Lemma 2.32 *Let G be a wallpaper group that contains an odd isometry with axis m . Then there exist translations $\tau_1, \tau_2 \in G$ that generate all translations of G and either*

- (1) τ_1 is parallel to m and τ_2 is perpendicular to m , or
- (2) τ_1 and τ_2 are of equal length and m is parallel to $\tau_1 \tau_2$.

Moreover, in case (2), group G contains a reflection.

In case (1) the translation lattice is rectangular, and m is parallel to a side of the rectangles, and in case (2) the translation lattice is rhombic, and m is parallel to a diagonal of the rhombi:



Proof. Let τ_A be the shortest translation in G , and let $\tau_B \in G$ be the shortest translation not generated by τ_A . According to Lemma 2.30, τ_A and τ_B generate all translations of G . Let $\alpha \in G$ be an odd isometry with axis m , that is, α is a glide reflection with axis m . Notice that for every translation τ

$$\alpha\tau\alpha^{-1} = \sigma_m\tau\sigma_m.$$

This follows from the facts that $\alpha = \sigma_m\tau'$ where τ' is a translation, and that translations commute.

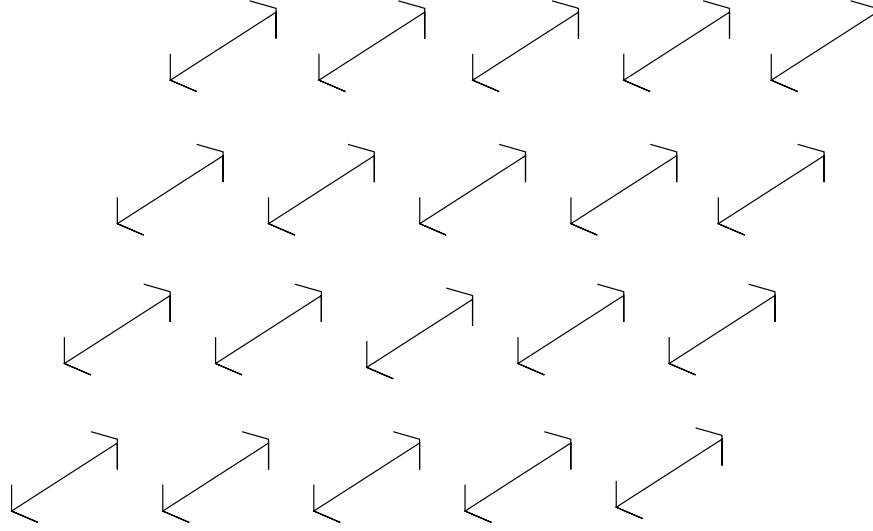
Consider the translation $\tau_C = \alpha\tau_A\alpha^{-1} = \sigma_m\tau_A\sigma_m$. It has the same length as the shortest translation τ_A . If τ_C is not generated by τ_A then it is the shortest translation not generated by τ_A , and according to Lemma 2.30 translations τ_A and τ_C generate all translations of G . If we choose $\tau_1 = \tau_A$ and $\tau_2 = \tau_C$ we have generating translations that satisfy the condition (2) of the lemma.

Assume then that τ_C is generated by τ_A . Then either $C = A$, in which case m is parallel to A , or $C = -A$, in which case m is perpendicular to A . Consider the conjugate $\tau_D = \alpha\tau_B\alpha^{-1}$ (if m is parallel to A) or $\tau_D = \alpha\tau_{-B}\alpha^{-1}$ (if m is perpendicular to A). In either case, $B + D$ is parallel to A . If $|B + D| > |A|$ then $B - A$ or $B + A$ is shorter than B , which contradicts the minimality of vector B . We must have $B + D = 0$ or $B + D = \pm A$. If $B + D = 0$ then B is perpendicular to A and we can choose $\tau_1 = \tau_A$, $\tau_2 = \tau_B$ and condition (1) of the lemma is satisfied. And if $B + D = \pm A$ then we choose $\tau_1 = \tau_B$, $\tau_2 = \tau_D$ (if m is parallel to A) or $\tau_1 = \tau_{-B}$, $\tau_2 = \tau_D$ (if m is perpendicular to A). In either case, condition (2) of the lemma is satisfied. Notice that τ_B and τ_D generate all translations because they generate τ_A .

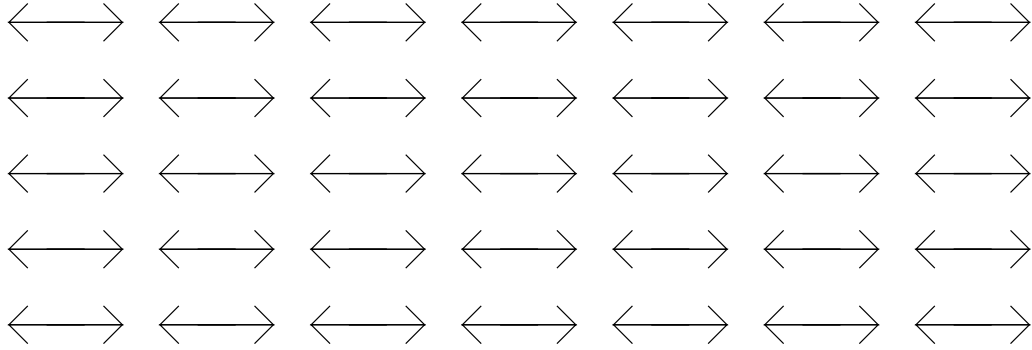
Finally, to prove the last claim, assume that case (2) applies. Because α^2 is a translation that is parallel to $\tau_1\tau_2$, we must have that $\alpha^2 = (\tau_1\tau_2)^i = \tau_2^i\tau_1^i$ for some integer i . Since translations τ_1 and τ_2 are conjugate by $\tau_2 = \alpha\tau_1\alpha^{-1}$, we also have that $\tau_2^i = \alpha\tau_1^i\alpha^{-1}$. This means that $\alpha^2 = \alpha\tau_1^i\alpha^{-1}\tau_1^i$. Divide both sides by α^2 from the left, and we have the result that $\alpha^{-1}\tau_1^i$ is an odd involution, that is, a reflection. $\square$

Our lemma limits the number of possible odd isometries of wallpaper groups sufficiently so that we can proceed with the analysis of the wallpaper groups G with halfturns and some odd isometries.

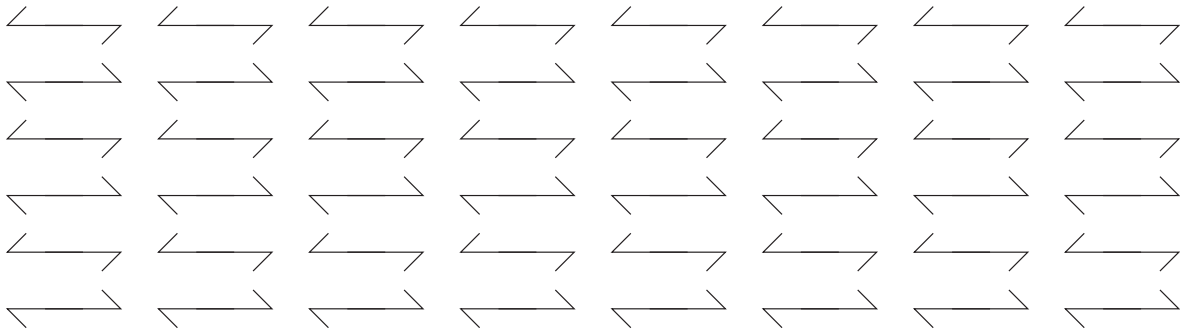
(a) First, assume that G contains a reflection σ_m such that the condition (2) of the previous lemma is satisfied. The lattice determined by the two generating translations from the lemma is rhombic. Let us prove that line m must contain a center of a halfturn. Consider a rhombus that is intersected by m , whose corners are centers of halfturns and whose interior does not contain any such centers. We know that m is parallel to a diagonal of the rhombus. If m is not the diagonal then one of the corners is mapped inside the rhombus by reflection σ_m , which contradicts the fact that there are no rotation centers inside the rhombus. We conclude that m bisects the rhombus along its diagonal, and therefore m contains a center of rotation. As the first halfturn σ_P was chosen arbitrarily, we can choose it in such a way that $P \in m$. We see that line m is then uniquely determined by τ_1, τ_2 and P . All other odd elements of G are then the products of σ_m and even isometries. This gives the wallpaper group $W_2^1 = \langle \tau_1, \tau_2, \sigma_P, \sigma_m \rangle$ where τ_1 and τ_2 are of equal length, and m is the line through P and $\tau_1\tau_2(P)$. This group is the symmetry group of



(b) Assume then that G contains a reflection σ_m that satisfies the condition (1) of the lemma. Let us call the direction of τ_1 and m the horizontal direction. We have two possibilities: (i) that m contains a center of a halfturn, and (ii) that m does not contain a center of a halfturn. In the second case the line m must run in the middle between two horizontal rows of rotation centers. As before, all other odd isometries are uniquely determined by σ_m and the even isometries. We get two wallpaper groups $W_2^2 = \langle \tau_1, \tau_2, \sigma_P, \sigma_m \rangle$ where m is the line through P and $\tau_1(P)$, and $W_2^3 = \langle \tau_1, \tau_2, \sigma_P, \sigma_m \rangle$ where m is the perpendicular bisector between points P and the center of halfturn $\tau_2\sigma_P$. In both cases, τ_1 and τ_2 are perpendicular. Group W_2^2 is the symmetry group of

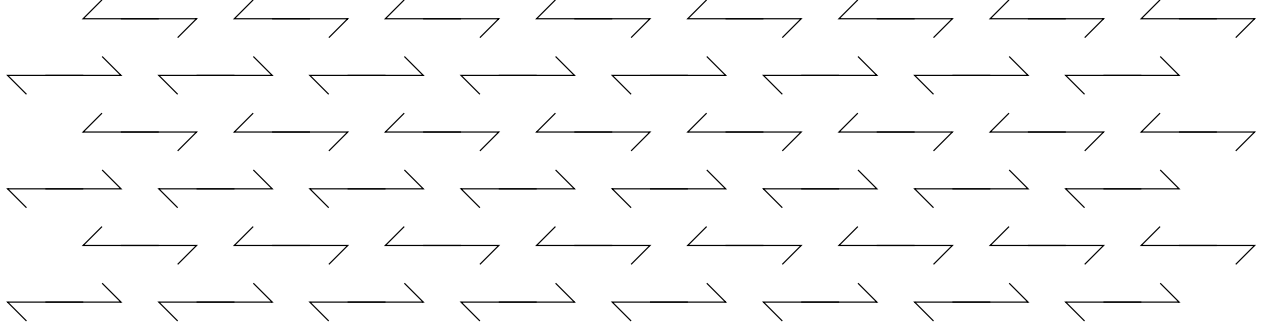


and group W_2^3 is the symmetry group of

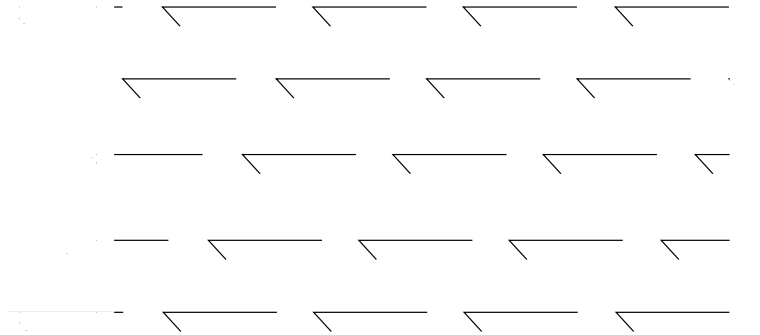


(c) Finally, assume that G does not contain any reflections. Let $\gamma \in G$ be a glide reflection with axis m . According to the last claim of Lemma 2.32, case (1) of the lemma must apply. Let τ_1 and τ_2 be two

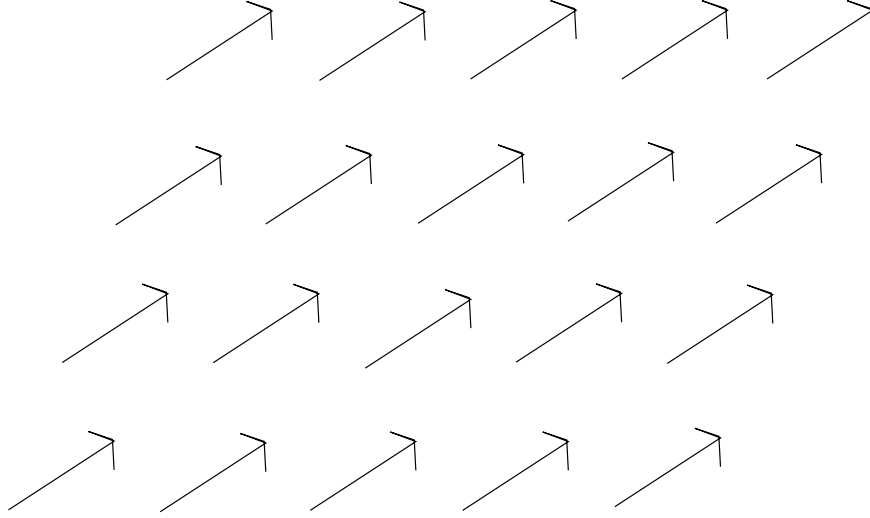
perpendicular translations, as indicated by the case (1) of the lemma. If the axis m contains the center P of some halfturn $\sigma_P \in G$ then G contains the reflection $\gamma\sigma_P$. We conclude that m must run in the middle between two horizontal rows of rotation centers. Let integer i be such that $\gamma^2 = \tau_1^i$. If i would be even then γ and τ_1 would generate a reflection, so i must be odd. By multiplying γ with a suitable power of τ_1 we obtain a glide reflection whose square is exactly τ_1 . This is uniquely determined, so the group G is also determined. It is $W_2^4 = \langle \tau_2, \sigma_P, \gamma \rangle$ where γ is a glide reflection such that $\tau_1 = \gamma^2$ and τ_2 are perpendicular. This is the symmetry group of



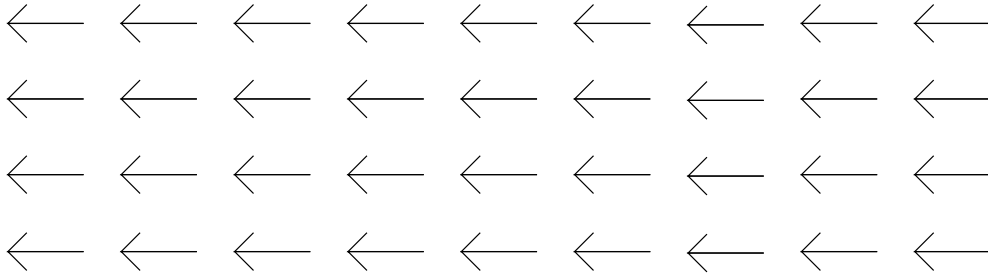
5) As our final case, assume that there are no non-trivial rotations in group G . The even isometries are then all translations generated by τ_1 and τ_2 . If there are no odd isometries then the group is $W_1 = \langle \tau_1, \tau_2 \rangle$. This group is the symmetry group of



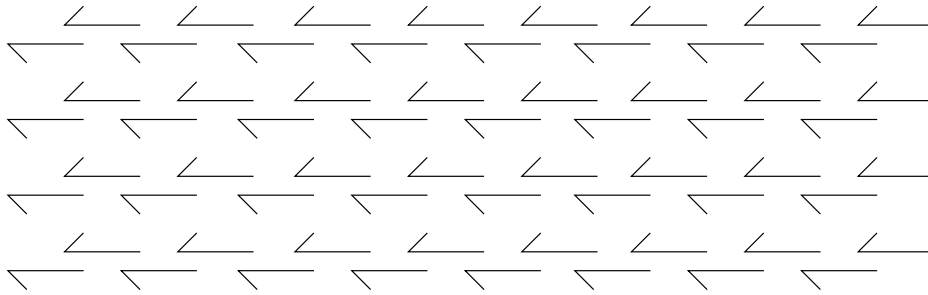
Let us assume then that G also contains odd isometries. If G contains a reflection σ_m then according to Lemma 2.32 either G has perpendicular generating translations τ_1 and τ_2 and m is parallel to τ_1 , or G has generating translations τ_1 and τ_2 of equal length and m is parallel to $\tau_1\tau_2$. In the second case we obtain group $W_1^1 = \langle \tau_1, \tau_2, \sigma_m \rangle$ that is the symmetry group of



and in the first case we obtain the symmetry group $W_1^2 = \langle \tau_1, \tau_2, \sigma_m \rangle$ of



Assume then that G does not contain any reflections but contains a glide reflection with axis m . Case (1) of Lemma 2.32 must apply. Then we can choose the glide reflection γ in such a way that $\tau_1 = \gamma^2$. This gives the last wallpaper group $W_1^3 = \langle \gamma, \tau_2 \rangle$. A pattern with this symmetry group is for example



We have exhausted all possibilities of wallpaper groups. We found 17 groups: Two with 6-fold rotations, three with 4-fold rotations, three with 3-fold (but no 6-fold) rotations, five with halfturns (but no higher order rotations) and four without non-trivial rotations.

Theorem 2.33 *Let G be a wallpaper group. Then G is among the 17 groups discussed above.* □

2.11 Final remarks on discrete symmetry groups

The rosette groups, frieze groups and the wallpaper groups have standard names given by crystallographers, and standardized by the International Union of Crystallography. Another naming system was developed by Fejes Tóth. The following table summarizes these notations:

Our notation	Fejes Tóth	Crystallographic
C_n	C_n	n
D_n	D_n	nm , if m is odd, nmm , if m is even
F_{1001}	F_1^1	p1m1
F_{1111}	F_2^1	pmm2
F_{0000}	F_1	p111
F_{0100}	F_1^2	pm11
F_{0010}	F_2	p112
F_{0001}	F_1^3	p1a1
F_{0111}	F_2^2	pma2
W_6	W_6	p6
W_6^1	W_6^1	p6m
W_3	W_3	p3
W_3^1	W_3^1	p3m1
W_3^2	W_3^2	p31m
W_4	W_4	p4
W_4^1	W_4^1	p4m
W_4^2	W_4^2	p4g
W_2	W_2	p2
W_2^1	W_2^1	cmm
W_2^2	W_2^2	pmm
W_2^3	W_2^3	pmg
W_2^4	W_2^4	pgg
W_1	W_1	p1
W_1^1	W_1^1	cm
W_1^2	W_1^2	pm
W_1^3	W_1^3	pg

Observe that each rosette, frieze or wallpaper group type is actually a family of subgroups of $\mathcal{I}$. For example, for each $P \in \mathbb{R}^2$, the halfturn around point P generates the cyclic group C_2 , but of course each choice of P provides a distinct subgroup of $\mathcal{I}$. In fact, each group type represents a family of affinely conjugate subgroups, as explained briefly below:

- An affine transformation of the plane is a transformation that preserves parallelism of lines. It is the composition of a linear transformation and a translation, that is, a mapping

$$f : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto M \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}$$

where M is a 2×2 matrix. The transformation is one-to-one if and only if M is invertible, i.e., $\det(M) \neq 0$. Isometries are exactly the distance preserving affine maps. Distance preservation is equivalent to M being an orthogonal matrix, i.e., equivalent to $MM^T = I$ where M^T is the transpose of M and I is the 2×2 identity matrix. Even and odd isometries correspond to orthogonal matrices M whose determinant is $+1$ and -1 , respectively.

- Two subgroups G_1 and G_2 of $\mathcal{I}$ are said to be equal up to affine conjugacy if there exists a one-to-one affine transformation f such that $G_1 = fG_2f^{-1}$, that is, elements of G_1 are exactly the functions $f\alpha f^{-1}$ for $\alpha \in G_2$. (In particular, this requires that $f\alpha f^{-1}$ are isometries for all $\alpha \in G_2$, which is not the case for all affine f and all isometries $\alpha \in \mathcal{I}$.)
- If G_1 and G_2 are wallpaper groups, frieze groups or rosette groups then equality up to affine conjugacy exactly means that they are of the same wallpaper, frieze or rosette group type.
- Affine conjugacy preserves isometry types: If α and $f\alpha f^{-1}$ are both isometries then they are of the same type: both translations, both rotations, both reflections or both glide reflections. (To see this, note that the parity of the isometry is preserved by affine conjugacy, and that P is a fixed point of α if and only if $f(P)$ is a fixed point of $f\alpha f^{-1}$.) But as mentioned above, $f\alpha f^{-1}$ may also not be an isometry.
- As groups, C_2 and D_1 are isomorphic. But they are not equal up to affine conjugacy. Likewise, frieze groups F_{0000} and F_{0001} are isomorphic (both are infinite cyclic groups, one is generated by a translation the other one by a glide reflection) but we consider them different as they are not affinely conjugate.

3 Tilings

Intuitively, a tiling is a covering of the plane without overlaps using some tiles. We start by giving more precise definitions. You may want to review some basic concepts of topology (especially the standard Euclidean topology of $\mathbb{R}^2$) such as

- open and closed sets,
- neighborhood of a point (=any open set containing the point),
- interior of a set (=largest open set contained in the set),
- closure of a set (=smallest closed set containing the set),
- boundary of a set (=intersection of the closures of the set and its complement),
- compactness,
- continuity of functions (inverse images of open sets are open),
- homeomorphism (=continuous bijection whose inverse is also continuous).
- connectedness (a set is connected iff it is not the union of two disjoint open sets),

Recall that since the Euclidean topology of $\mathbb{R}^2$ is metric, it is Hausdorff, and compactness is equivalent to being closed and bounded. Also, in $\mathbb{R}^2$ an open set is connected if and only if it is path-connected, that is, each pair of its points can be joined by a path (=homeomorphic image of the unit interval) inside the set. Let us denote by

$$B_r(P) = \{X \in \mathbb{R}^2 \mid d(X, P) < r\}$$

the open disk of radius r centered at P , and if P is the origin O , we simply denote $B_r = B_r(O)$. The closure of an open disk is a closed disk

$$\overline{B}_r(P) = \{X \in \mathbb{R}^2 \mid d(X, P) \leq r\},$$

and $\overline{B}_r = \overline{B}_r(O)$.

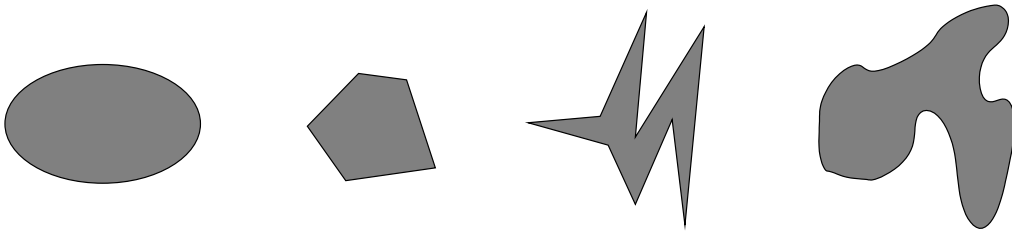
3.1 Basic definitions

A tile is a subset of $\mathbb{R}^2$ that is a topological disk. This means that it is the image of the closed disk $\overline{B}_1$ under some homeomorphism. Homeomorphisms preserve topological properties, so tile t immediately inherits topological properties from the disk $\overline{B}_1$:

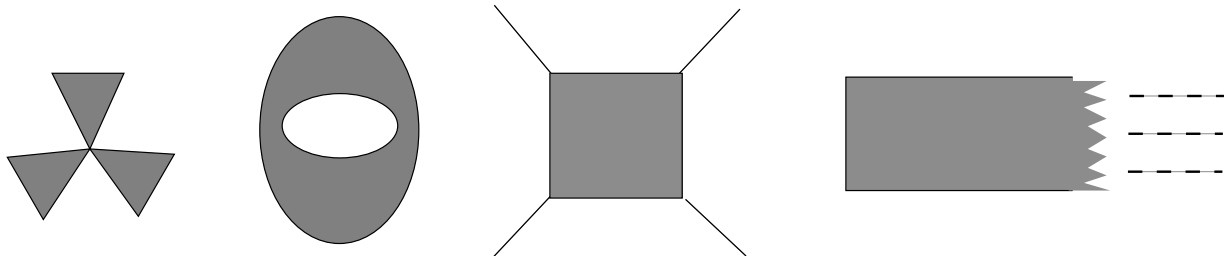
- t is compact (=closed and bounded),
- the interior of t is connected, and the complement of t is connected,
- the boundary of t is the boundary of its interior,
- the boundary of t is a simple closed curve, that is, homeomorphic to the unit circle

$$\{X \in \mathbb{R}^2 \mid d(X, O) = 1\}.$$

This definition of a tile is very general. Later, additional restrictions will be added as needed. For example, we may restrict our attention to tiles that are polygons. Here are some examples of tiles:



but these are not tiles:



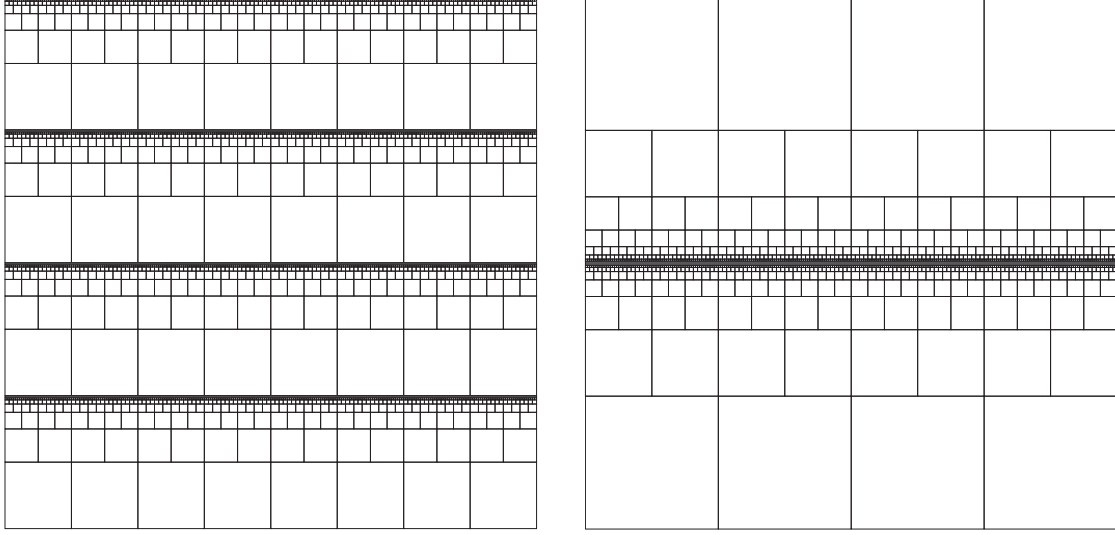
(They are with non-connected interior, non-connected complement, boundary that is not the boundary of the interior, and unbounded, in this order.)

A tiling $\mathcal{T}$ is a family of tiles that covers the plane

- (1) without gaps (every $P \in \mathbb{R}^2$ belongs at least one tile), and
- (2) without overlaps (the interiors of the tiles are pairwise disjoint).

Notice that the boundaries of the tiles do not need to be disjoint. But it follows that every point that belongs to more than one tile cannot belong to the interior of any tile. Notice also that the number of tiles in any tiling must be infinite (union of a finite number of bounded sets would be bounded) but countable (the interior of each tile contains a point with rational coordinates).

This definition of tilings is very general. It does not restrict the number of different shapes used in any way, so one tiling can, for example, contain arbitrarily small tiles. The left picture below represents a part of a tiling, while the rightmost picture is not a tiling since the horizontal line in the center is not covered by any tile.



Let $\mathcal{T} = \{t_1, t_2, \dots\}$ be a tiling. Its symmetry group G consists of those isometries α that take every tile of $\mathcal{T}$ onto a tile of $\mathcal{T}$, that is, for every $i = 1, 2, \dots$ there exists j such that $\alpha(t_i) = t_j$. It is easy to see that symmetry groups of tilings (even under our very general definition of tiles) are discrete: the only possibilities are our familiar rosette, frieze and wallpaper groups.

Theorem 3.1 *The symmetry group of a tiling is discrete.*

Proof. Let G be the symmetry group of tiling $\mathcal{T} = \{t_1, t_2, \dots\}$. Then there must exist a positive number ε such that the length of every non-trivial translation in G is at least ε . Indeed, the interior of tile t_1 contains a disk $B_\varepsilon(P)$ for some $\varepsilon > 0$, so any translation τ that is shorter than ε takes P into the interior of t_1 . This means that $\tau(t_1) = t_1$, which is possible only if $\tau = \text{id}$.

Consider then rotations. Suppose first there is a non-trivial translation τ in G . If there are arbitrarily small rotations in G then there are arbitrarily small translations among $\tau^{-1}\rho\tau\rho^{-1}$, which contradicts the conclusion in the previous paragraph.

Suppose then that G contains only the trivial translation. Then all rotations have the same center P of rotation (Corollary 2.22). Suppose there would be arbitrarily small rotations around P .

Let $t \in \mathcal{T}$ be a tile that contains point P . We have $t \subseteq B_k(P)$ for a sufficiently large number k . Let Q be a point whose distance from P is at least k such that Q belongs to the interior of some tile $t' \in \mathcal{T}$. (Just choose any point Q sufficiently far away from P . If Q is not in the interior of any tile then Q is on the boundary of some t' . There are interior points of t' close to Q . We can choose any one of them.)

The circle $c = \{X \in \mathbb{R}^2 \mid d(P, X) = d(P, Q)\}$ does not intersect t , but it contains an interior point Q of t' . Let us prove that $c \subseteq t'$. Assume the contrary: there exists a point $R \in c$ such that $R \notin t'$. The complement of t' is open so, for all sufficiently small angles Θ , we have $\rho_{P, \Theta}(R) \notin t'$.

Let $\varepsilon > 0$ be a small number so that $\rho_{P, \Theta}(Q)$ is an interior point of t' and $\rho_{P, \Theta}(R) \notin t'$ for all angles Θ with $|\Theta| < \varepsilon$. Choose one positive angle $\Theta < \varepsilon$ such that $\rho = \rho_{P, \Theta} \in G$. Because ρ is a symmetry of the tiling such that $\rho(Q)$ is an interior point of t' , we must have that $\rho(t') = t'$. This means that $\rho^i(Q) \in t'$ for all integers i . Choose number i such that $|i\Theta - \Phi| < \varepsilon$ where Φ is the angle such that $\rho_{P, \Phi}(Q) = R$. Then $\rho^i(Q) \in t'$ but, on the other hand,

$$\rho^i(Q) = \rho_{P, i\Theta}(Q) = \rho_{P, i\Theta - \Phi} \rho_{P, \Phi}(Q) = \rho_{P, i\Theta - \Phi}(R) \notin t',$$

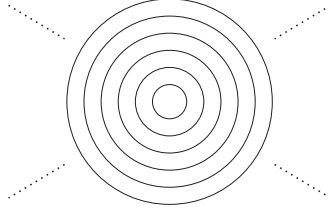
a contradiction.

We have proved that $c \subseteq t'$. Then the complement of t' is not connected: Interior points of t are in the disk $B_k(P)$ so they are separated by t' from the points outside the circle c . This contradicts the fact

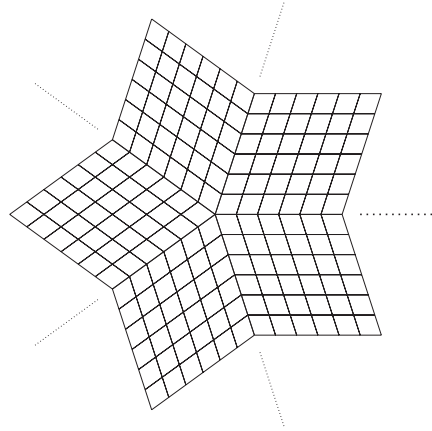
that t' should be a topological disk. Conclusion: there can only be a finite number of rotations in G , so G is a finite subgroup of $\mathcal{I}$, and therefore a rosette group.

□

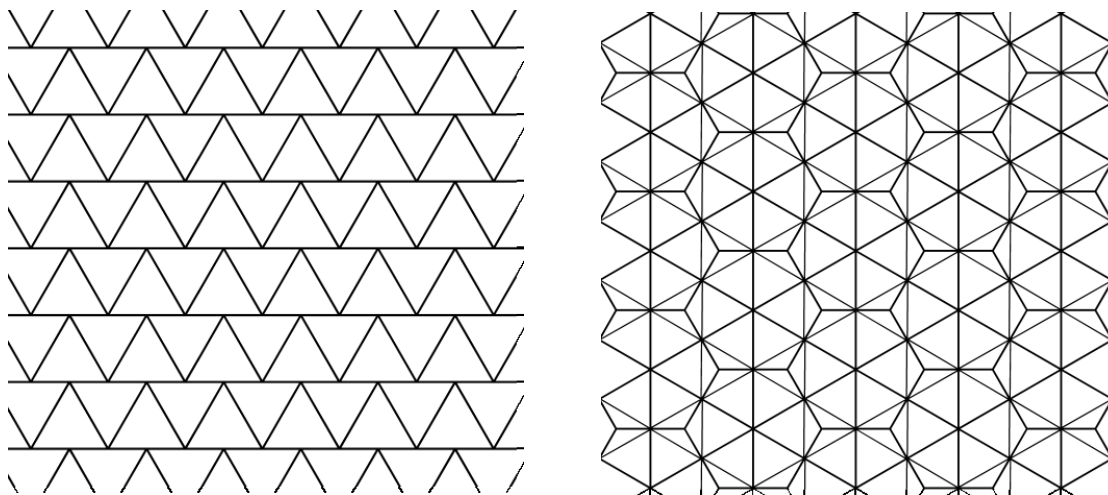
Note that it is essential in the proof that the tiles are topological disks, and hence do not contain holes. If we would allow tiles that are topological rings then we would have, for example, the following "tiling" whose symmetry group is not discrete.



Each rosette group, frieze group and wallpaper group is the symmetry group of some tiling. We see some examples in the homeworks. As another example, below is a piece of a tiling whose symmetry group is D_5 . This can be easily generalized to obtain a tiling whose symmetry group is D_n or C_n , for any $n \geq 5$.



Our main interest is in tilings using only a finite number of different shapes. More precisely, tiles $\{p_1, p_2, \dots, p_k\}$ are prototiles of a tiling $\mathcal{T} = \{t_1, t_2, \dots\}$ if every tile $t_i \in \mathcal{T}$ is congruent to some p_j . By congruent we mean that there is an isometry (even or odd!) that takes t_i onto p_j . We say that the prototiles $\{p_1, p_2, \dots, p_k\}$ admit the tiling $\mathcal{T}$. Tiling $\mathcal{T}$ is called k -hedral, where k is the number of prototiles p_j . In the special cases of $k = 1$ and $k = 2$ the tiling is called monohedral and dihedral, respectively. Note that some tiles may be "flipped over" copies of the prototiles, that is, the isometry that takes the prototile on a tile may be odd. In some cases we may be interested in those k -hedral tilings where the tiles are congruent to prototiles by even isometries, but in these cases this will be stated explicitly. Here is an example of a monohedral and a dihedral tiling:

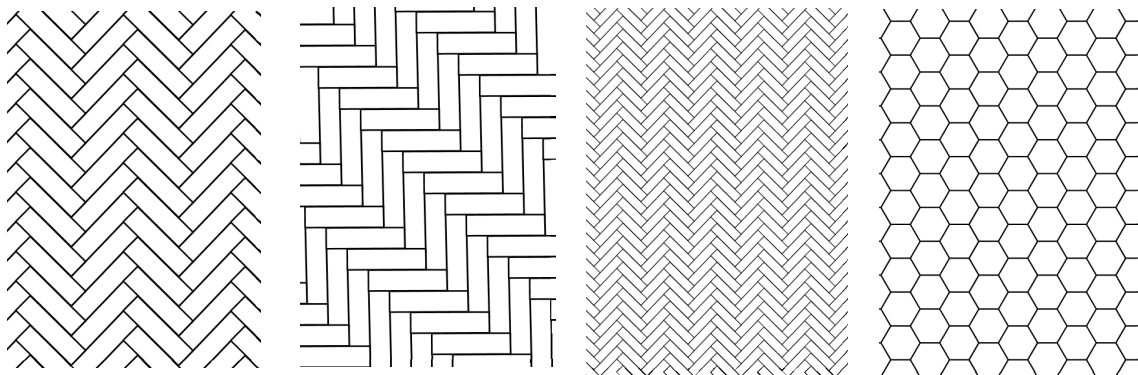


Let $\mathcal{T} = \{t_1, t_2, t_3, \dots\}$ be a tiling. If $h : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a homeomorphism then also $h(\mathcal{T}) = \{h(t_1), h(t_2), h(t_3), \dots\}$ is a tiling. We say that tilings $\mathcal{T}$ and $h(\mathcal{T})$ are topologically equivalent. This is easily seen to be an equivalence relation among tilings.

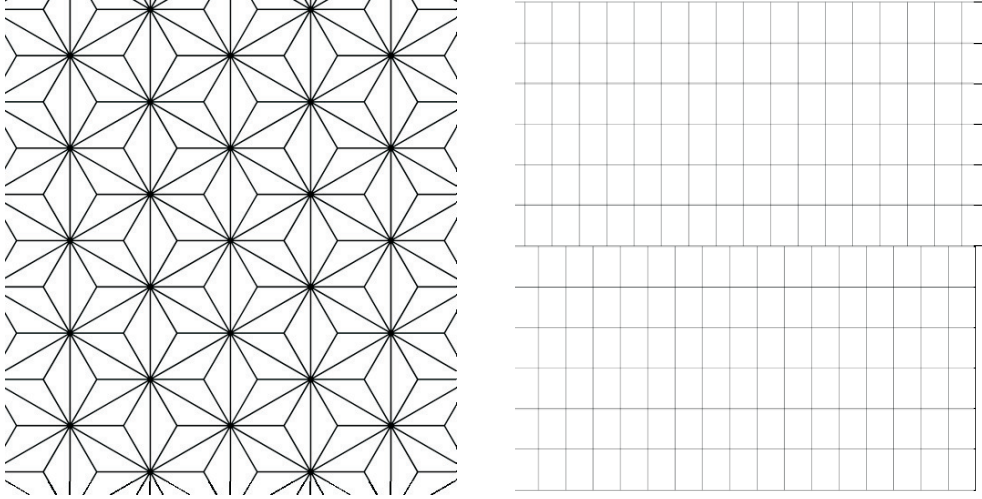
Every isometry is a homeomorphism, so if α is an isometry then $\alpha(\mathcal{T}) = \{\alpha(t_1), \alpha(t_2), \alpha(t_3), \dots\}$ is a tiling. We say that that $\alpha(\mathcal{T})$ is congruent to tiling $\mathcal{T}$. Also congruence is an equivalence relation among tilings.

Finally, a similarity $s : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a composition of an isometry and a stretch (that is, a function that maps $(x, y) \mapsto (kx, ky)$ for some $k > 0$). In other words, a similarity s by factor $k > 0$ is a function such that for any two points $P, Q \in \mathbb{R}^2$ we have $d(s(P), s(Q)) = k \cdot d(P, Q)$. Similarities are homeomorphisms, so $s(\mathcal{T}) = \{s(t_1), s(t_2), s(t_3), \dots\}$ is a tiling. We say that tilings $\mathcal{T}$ and $s(\mathcal{T})$ are similar. Intuitively, similarity of two tiling means that they look the same when one of them is watched under a suitable magnifying class. Usually (unless otherwise noted) we consider similar tilings to be the same tiling.

The following figure contains four topologically equivalent monohedral tilings. First two are congruent with each other, and they are similar to the third one:



Two tiles t_1 and t_2 of tiling $\mathcal{T}$ are called equivalent in $\mathcal{T}$ if there exists a symmetry of $\mathcal{T}$ that takes t_1 onto t_2 . This is clearly an equivalence relation among tiles t_i . Equivalence classes are called the transitivity classes of $\mathcal{T}$. If tiling $\mathcal{T}$ has only one transitivity class then the tiling is called isohedral (or tile-transitive). More generally, if there are k transitivity classes then the tiling is called k -isohedral. Notice that any isohedral tiling is monohedral as equivalent tiles are congruent. But there are monohedral tilings that are not isohedral. Analogously, a k -isohedral tiling is always k -hedral (but it can also be n -hedral for some $n < k$). Here are examples of an isohedral tiling and a monohedral tiling that is not isohedral, or even k -isohedral for any finite k .

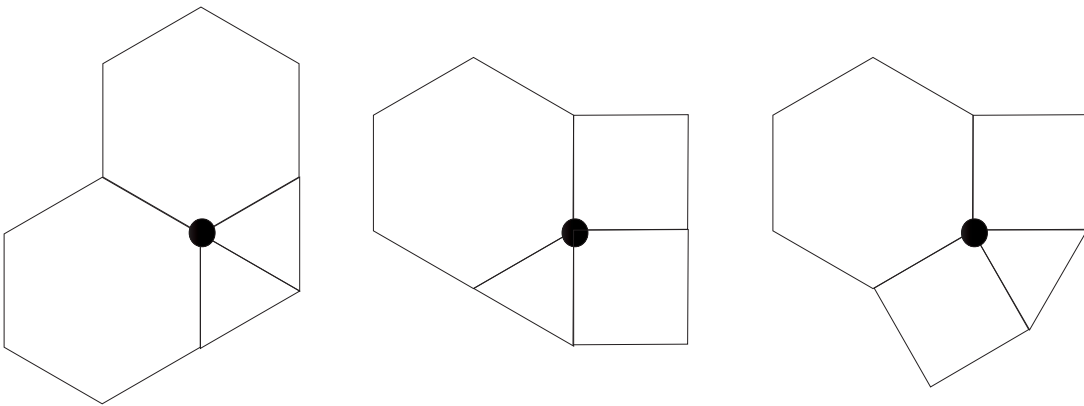


It is easy to see (in the homeworks!) that the symmetry group of a k -hedral tiling is a wallpaper group if and only if the tiling is n -isohedral for some n . (But there are also tilings that are not k -hedral for any k and whose symmetry group is a wallpaper group.)

3.2 Tilings by regular polygons

We restrict the study in this section to tilings that are by regular polygons, and that are edge-to-edge, that is, the intersection of two tiles is either empty, single vertex of the polygons, or the entire edge of the two neighboring polygons. Two tiles are called edge neighbors (vertex neighbors) if their intersection is an edge (edge or vertex, respectively) of the polygons. Corners of the polygons are called the vertices of the tiling.

Consider a vertex P where r regular polygons of orders $n_1, n_2, n_3, \dots, n_r$ meet, in this order (counted clockwise or counterclockwise). Then we say that the vertex is of type $n_1 \cdot n_2 \cdot \dots \cdot n_r$. For example, vertices of types $3 \cdot 3 \cdot 6 \cdot 6$, $3 \cdot 4 \cdot 4 \cdot 6$ and $3 \cdot 4 \cdot 6 \cdot 4$ look like



Notice that types $3 \cdot 4 \cdot 4 \cdot 6$ and $4 \cdot 6 \cdot 3 \cdot 4$ and $4 \cdot 3 \cdot 6 \cdot 4$ are all identical, as they are obtained by changing the starting point and/or the direction of reading the polygons. We also adapt the usual shorthand notations for repetitions, so that $3 \cdot 3 \cdot 6 \cdot 6$ may be abbreviated as $3^2 \cdot 6^2$.

The interior angle of a regular n -gon is $180^\circ(1 - \frac{2}{n})$. Consequently, if P is a vertex of type $n_1 \cdot n_2 \cdot \dots \cdot n_r$ then

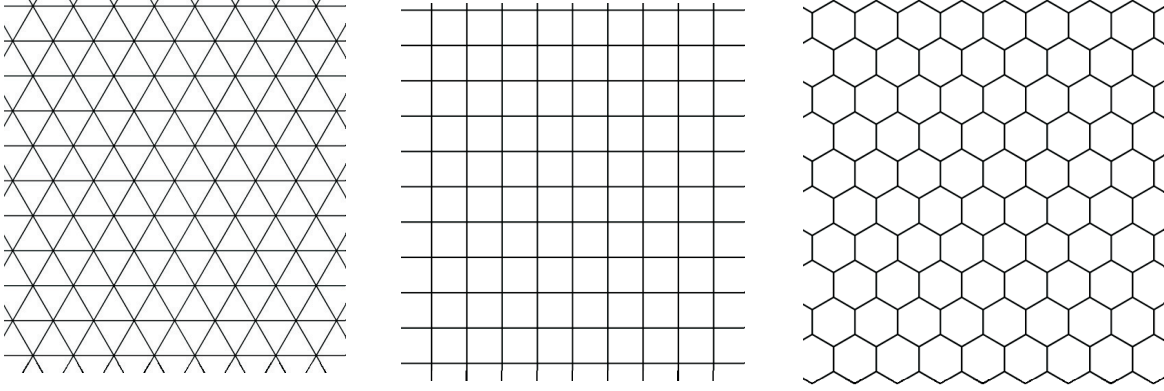
$$\sum_{i=1}^r \left(1 - \frac{2}{n_i}\right) = 2. \quad (1)$$

This follows from the fact that the interior angles of the polygons that meet at P must sum up to 360° .

Assume first that the tiling is monohedral, with all tiles regular n -gons. Then (1) becomes

$$r(1 - \frac{2}{n}) = 2,$$

which implies $n = \frac{2r}{r-2}$. Because n is positive, we must have $r \geq 3$, and because $n \geq 3$ we must have $r \leq 6$. With $r = 3, 4, 5$ and 6 we get $n = 6, 4, \frac{10}{3}$ and 3 . Number n is an integer so we only have three solutions. These are the familiar regular tilings



Theorem 3.2 *The only edge-to-edge monohedral tilings by regular polygons are the three regular tilings above.* □

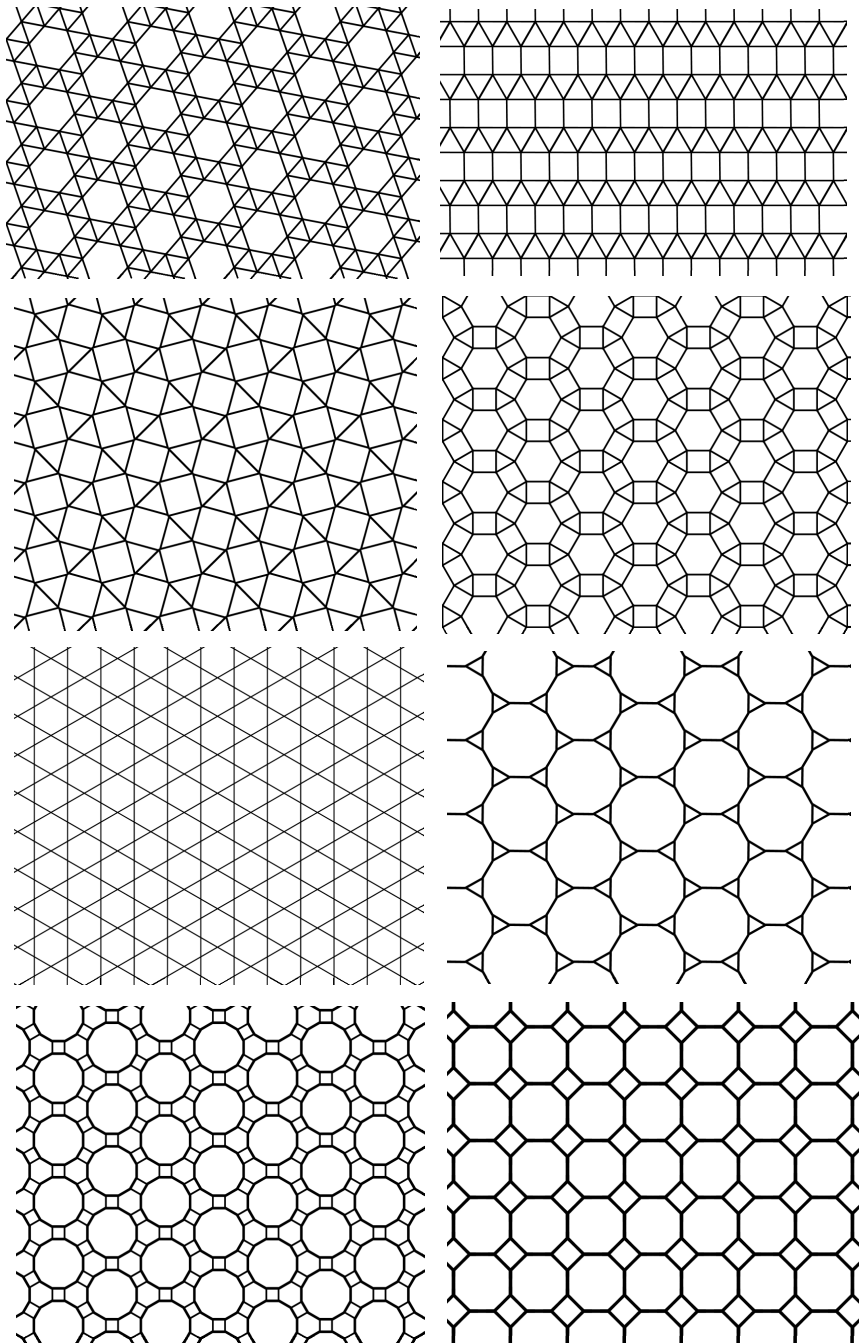
Consider then the case when the tiling is not necessarily monohedral. Possible types of vertices are limited by (1). We only have the following numerical solutions to (1), and the corresponding possibilities for the vertex types:

type	archimedean
$3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$	A
$3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$	A
$3 \cdot 3 \cdot 3 \cdot 4 \cdot 4$	A
$3 \cdot 3 \cdot 4 \cdot 3 \cdot 4$	A
$3 \cdot 3 \cdot 4 \cdot 12$	
$3 \cdot 3 \cdot 6 \cdot 6$	
$3 \cdot 4 \cdot 3 \cdot 12$	
$3 \cdot 4 \cdot 4 \cdot 6$	
$3 \cdot 4 \cdot 6 \cdot 4$	A
$3 \cdot 6 \cdot 3 \cdot 6$	A
$3 \cdot 7 \cdot 42$	
$3 \cdot 8 \cdot 24$	
$3 \cdot 9 \cdot 18$	
$3 \cdot 10 \cdot 15$	
$3 \cdot 12 \cdot 12$	A
$4 \cdot 4 \cdot 4 \cdot 4$	A
$4 \cdot 5 \cdot 20$	
$4 \cdot 6 \cdot 12$	A
$4 \cdot 8 \cdot 8$	A
$5 \cdot 5 \cdot 10$	
$6 \cdot 6 \cdot 6$	A

The last column indicates whether the vertex type appears in some archimedean tiling: An edge-to-edge tiling by regular polygons is termed archimedean if all vertices of the tiling are of the same type. The three regular tilings are all archimedean, corresponding to vertex types 6^3 , 4^4 and 3^6 . In addition, it turns out that there are only eight other examples of archimedean tilings, corresponding to the vertex types marked by "A" in the table above.

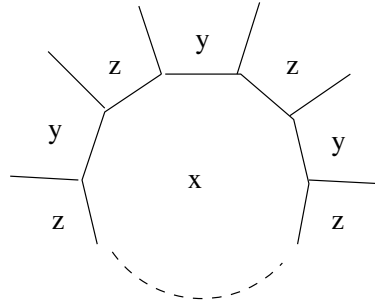
Theorem 3.3 (Kepler 1619) *There are exactly eleven different archimedean tilings, one of each type indicated by "A" in the table above.*

Proof. The eight non-regular archimedean tilings are shown below. It is easy to verify that they are indeed archimedean, and one can easily verify that the types of their vertices match the types marked by "A".



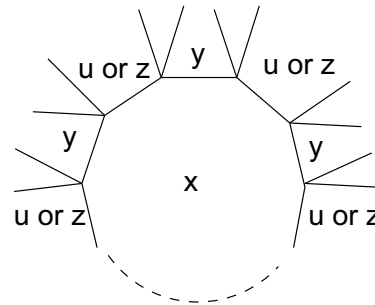
To prove that no other archimedean tilings exist we have to show that (i) the vertex types without "A" in the table are not possible in archimedean tilings, and (ii) each type with "A" leads to a unique tiling. Let us use the following terminology: a polygon is incident to its vertices and edges, and an edge is incident to its endpoints. Two vertices are adjacent if they are the two endpoints of an edge.

(i) Vertex type $x \cdot y \cdot z$ where x is odd and $y \neq z$ is not possible in any archimedean tiling: The edge neighbors of an x -gon across two consecutive edges are a y -gon and a z -gon. (Note: This is true even if $x = y$ or $x = z$.) So y -gons and z -gons alternate as the edge neighbors of an x -gon when we go around its edges clockwise. But since x is odd this is not possible: we necessarily end up with two consecutive neighbors of the same type.



This reasoning rules out six vertex types $3 \cdot 7 \cdot 42$, $3 \cdot 8 \cdot 24$, $3 \cdot 9 \cdot 18$, $3 \cdot 10 \cdot 15$, $4 \cdot 5 \cdot 20$ and $5 \cdot 5 \cdot 10$.

By a similar argument, vertex type $x \cdot y \cdot u \cdot z$ is not possible when x is odd, $y \neq z$, and no three of the numbers are equal. Clearly $x \neq y$ or $x \neq z$. The two situations are symmetric, so we may assume that $x \neq z$. Then two consecutive edge neighbors of an x -gon are an y -gon and a z -gon, or — if $x = y$ — possibly a y -gon and a u -gon. In either case, every other edge neighbor is a y -gon, and every other neighbor is not a y -gon, which is not possible as x is odd.

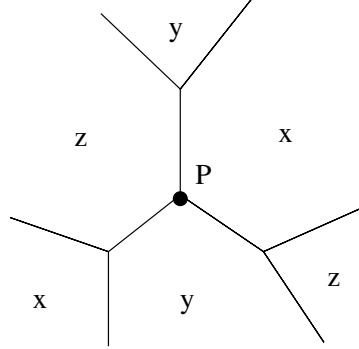


This rules out the remaining four vertex types $3 \cdot 3 \cdot 4 \cdot 12$, $3 \cdot 3 \cdot 6 \cdot 6$, $3 \cdot 4 \cdot 3 \cdot 12$ and $3 \cdot 4 \cdot 4 \cdot 6$.

(ii) Let us prove that any archimedean tiling $\mathcal{T}$ is similar to one of the given eleven tilings, namely the one with the same vertex type. We start by selecting one arbitrary vertex P of $\mathcal{T}$ and one arbitrary vertex P' of the known archimedean tiling $\mathcal{A}$ of the correct vertex type. There clearly exists a similarity function s that maps P onto P' in such a way that the polygons incident to P in $\mathcal{T}$ are mapped onto the polygons incident to P' in $\mathcal{A}$. Let us show that (with one exception in type $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$) similarity s maps the entire tiling $\mathcal{T}$ onto tiling $\mathcal{A}$.

It is enough to consider the vertices that are adjacent to P , and to show that all tiles incident to those vertices are mapped by s onto similar tiles on tiling $\mathcal{A}$. Namely then we can repeat the reasoning on the adjacent vertices to conclude that all vertices adjacent to them are mapped correctly, and so on, by mathematical induction, that all tiles at any distance from P are mapped onto tiles of $\mathcal{A}$.

Consider first tilings of vertex types $3 \cdot 12 \cdot 12$, $4 \cdot 6 \cdot 12$, $4 \cdot 8 \cdot 8$, and $6 \cdot 6 \cdot 6$, that is, the cases $x \cdot y \cdot z$ where three polygons meet at the vertices. Let Q be any of the three vertices adjacent to P . Two polygons incident to Q are also incident to P so they are known. This means that also the third polygon incident to Q is known and it must be mapped by s onto the corresponding tile in the archimedean tiling $\mathcal{A}$.

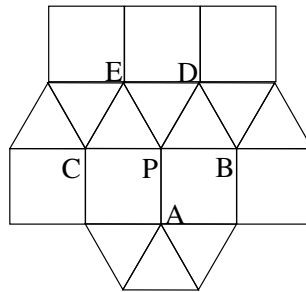


As discussed above, this is enough to prove that the entire tiling $\mathcal{T}$ is mapped onto $\mathcal{A}$.

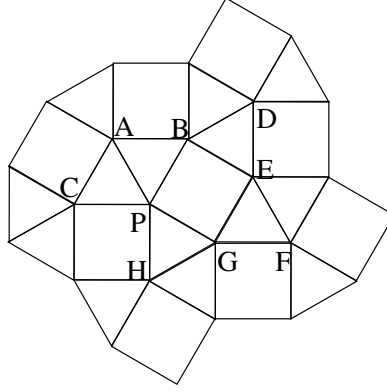
Vertex types $4 \cdot 4 \cdot 4 \cdot 4$ and $3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$ are also trivial: the polygons are all congruent and they must be correctly mapped onto the corresponding tiles in $\mathcal{A}$.

Consider then the vertex types $3 \cdot 4 \cdot 6 \cdot 4$ and $3 \cdot 6 \cdot 3 \cdot 6$. Let Q be a vertex adjacent to P . Two polygons that are edge neighbors and incident to Q are known. The other two are then also uniquely determined: in the first case one of the known polygons is a square, and the polygon opposite to it at Q must be a square as well, and in the case of $3 \cdot 6 \cdot 3 \cdot 6$ one of the known polygons is a triangle, and the polygon opposite to it at Q is a triangle. In both cases the polygons incident to Q are uniquely determined, and therefore mapped by s onto similar tiles in the tiling $\mathcal{A}$.

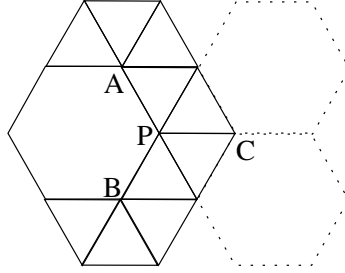
There remain three vertex types to analyze, namely $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$, $3 \cdot 3 \cdot 3 \cdot 4 \cdot 4$ and $3 \cdot 3 \cdot 4 \cdot 3 \cdot 4$. Consider type $3 \cdot 3 \cdot 3 \cdot 4 \cdot 4$ first: The following figure shows the order in which the vertices adjacent to P can be processed to determine the polygons incident to them. One can easily verify that the polygons are uniquely determined if the vertices are processed in the alphabetical order $A, B, C, D, \dots$. So the tiles are all mapped correctly onto tiling $\mathcal{A}$.



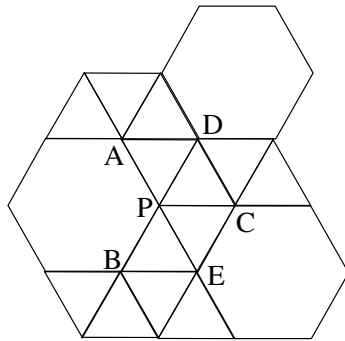
Analogously, if the vertex type is $3 \cdot 3 \cdot 4 \cdot 3 \cdot 4$ the vertices should be processed in the order indicated in this figure:



Finally, consider the vertex type $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$. In all previous cases, any similarity s that takes a vertex P and the incident polygons of $\mathcal{T}$ onto a vertex P' and its incident polygons of $\mathcal{A}$ is necessarily a similarity between entire tilings $\mathcal{T}$ and $\mathcal{A}$. But in the case of vertices of the type $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$ this is no longer true. Instead, there exist two similarities from vertex P onto vertex P' : one even and one odd similarity. And exactly one of them is a similarity between tilings $\mathcal{T}$ and $\mathcal{A}$. In the following figure, the polygons incident to vertices A and B are uniquely determined. Then, the hexagon incident to vertex C must be one of the two dotted hexagons in the illustration. (The third alternative would lead to two hexagons that are vertex neighbors, and is therefore impossible.)



In either case, the similarity s can be chosen in such a way that the hexagon incident to C is mapped correctly onto tiling $\mathcal{A}$. The similarity is even or odd depending on the position of the hexagon. Thereafter, the remaining polygons are uniquely determined. In this case we have to verify the uniqueness of the polygons up to vertices of distance two from P . After this the uniqueness of the entire tiling follows by mathematical induction:



□

Notice that the previous proof indicates that the 11 archimedean tilings are vertex transitive: for any two vertices P_1 and P_2 of the tiling, there exists a symmetry of the tiling that takes P_1 onto P_2 . With

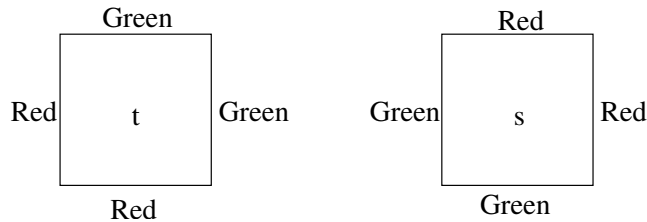
the exception of type $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$, any isometry that takes vertex P_1 and its incident polygons onto P_2 and its incident polygons is a symmetry of the tiling. The archimedean tiling of type $3 \cdot 3 \cdot 3 \cdot 3 \cdot 6$ comes in two enantiomorphic forms that are congruent with each other only by odd isometries.

Archimedean tilings are also called uniform, which refers to the fact that they are vertex transitive: the entire tiling looks exactly the same from each vertex. This is a stronger property than the property we started with: that the tiling looks locally the same at each vertex, as each vertex is of the same type.

4 Wang tiles

Wang tiles are unit square tiles with colored edges. Hence each tile can be represented as a 4-tuple (N, E, S, W) where N, E, S and W are the colors of the north, east, south and west sides of the square. Tilings with a finite number of prototiles are only considered. In Wang tilings copies of the prototiles are placed at integer lattice points, without rotating or flipping the tiles, so that all tiles are congruent to the given prototiles by translations only. A tiling can then be represented as a function $f : \mathbb{Z}^2 \rightarrow \mathcal{P}$ where $\mathcal{P}$ is the set of prototiles and $f(i, j)$ gives the tile at position $(i, j) \in \mathbb{Z}^2$. The tiling rule is that in a valid tiling the shared edge between any two tiles that are edge neighbors must have the same color.

For example, set $\mathcal{P} = \{(\text{Green}, \text{Green}, \text{Red}, \text{Red}), (\text{Red}, \text{Red}, \text{Green}, \text{Green})\}$ consists of two prototiles



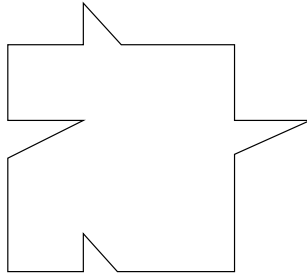
that admit the checkerboard-tiling

t	s	t	s	t	s	t
s	t	s	t	s	t	s
t	s	t	s	t	s	t
s	t	s	t	s	t	s

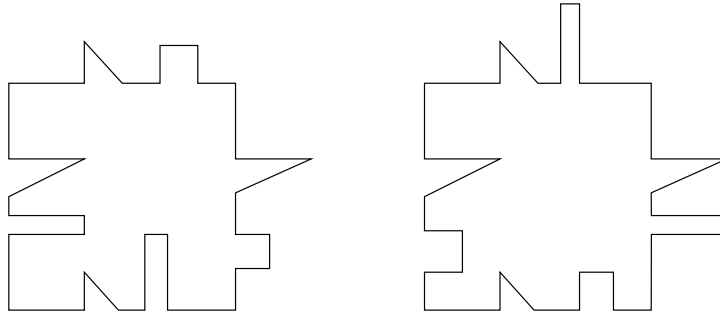
Wang tiles provide a discrete abstraction of tilings that allows us to study tilings using tools of discrete mathematics rather than geometry. This is especially useful when investigating computational properties and problems related to tilings. At first, Wang tiles may seem very restricted as the tilings are on a square lattice only. Nevertheless, the computational problems on Wang tiles are as hard as on more general types of tiles. By using Wang tiles we avoid problems related to representations of tiles (e.g. irrational coordinates of vertices) on computers, and we transform geometric problems into more manageable symbolic problems.

Our first observation is that Wang tiles fit our original definition of tiles as topological disks. We can namely represent Wang tiles as polygons as follows: The basic shape is a unit square. The middle of the north and east sides of each tile contain triangular "bumps" and the south and west sides have

”dents” that exactly fit the bumps. The bump/dent pairs are different in the horizontal and the vertical directions, and they are asymmetric so that flipped and non-flipped tiles do not match:



It should be clear that these tiles can only tile the plane in such a way that all tiles are aligned, and rotations and flips are not possible. To simulate the colors, we introduce an additional bump/dent pair on the sides of the tiles. Each color has its own bump/dent shape that does not fit with any other color. For example, our sample protoset of two tiles could look like this:



It should be obvious from this construction that any tiling by such polygons is congruent to a tiling where the tiles are positioned at integer lattice points, without rotations and flips. Such tilings are clearly ”isomorphic” to Wang tilings.

In the following we consider all tilings that given prototiles admit. In particular, we are interested to know when do given prototiles admit at least some tiling and when do they admit a periodic tiling. As it turns out that even among Wang tiles these questions can not be algorithmically answered (they are undecidable), so it follows that the questions are undecidable also among tiles that are polygons.

In the following two subsections we prove two preliminary results that will be needed in the algorithmic considerations that follow: First we show that if a finite set of Wang prototiles admits a tiling whose symmetry is a frieze group then it automatically admits also a tiling with a wallpaper symmetry. Then we prove that if one can tile arbitrarily large squares then one can also tile the entire infinite plane.

4.1 Periodic tilings

A tiling is called non-periodic if its symmetry group is finite, that is, if there is no translation that keeps the tiling invariant. A tiling is two-way periodic, or simply periodic, if its symmetry group is a wallpaper group, that is, if there are translations in non-parallel directions that keep the tiling invariant. A tiling whose symmetry group contains some non-trivial translation will be called one-way periodic.

Vector $(a, b) \neq (0, 0)$ is called a period of a tiling, if $\tau_{(a,b)}$ is a symmetry of the tiling. In the case of a Wang tiling $f : \mathbb{Z}^2 \rightarrow \mathcal{P}$ this means that $f(x, y) = f(x + a, y + b)$ for all $(x, y) \in \mathbb{Z}^2$.

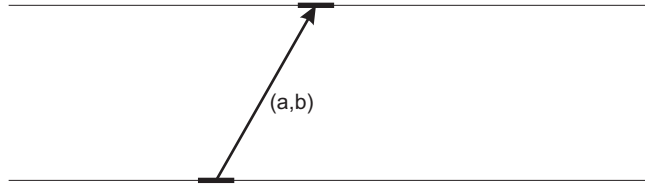
Notice that any two-way periodic tiling with Wang tiles has horizontal and vertical periods of equal lengths. Namely, if $f : \mathbb{Z}^2 \rightarrow \mathcal{P}$ is periodic with non-parallel periods (a, b) and (c, d) then it is also

periodic with the horizontal period $d(a, b) - b(c, d) = (ad - bc, 0)$ and the vertical period $a(c, d) - c(a, b) = (0, ad - bc)$. Note that $ad - bc \neq 0$ as vectors (a, b) and (c, d) are not parallel. In other words, a two-way periodic Wang tiling consists of a periodic repetition of a square pattern.

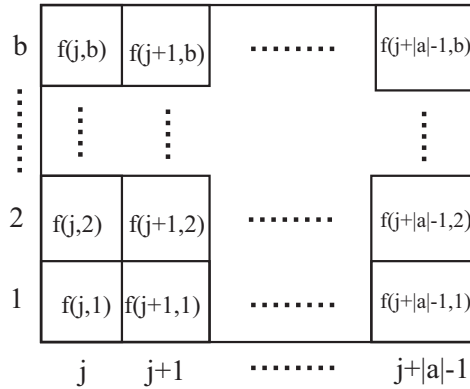
The next theorem states that a set of Wang tiles that admits a one-way periodic tiling also admits a periodic tiling:

Theorem 4.1 *Let $\mathcal{P}$ be a finite set of Wang prototiles that admits a tiling $f : \mathbb{Z}^2 \rightarrow \mathcal{P}$ that is one-way periodic. Then there exists also a two-way periodic tiling $g : \mathbb{Z}^2 \rightarrow \mathcal{P}$.*

Proof. Let $(a, b) \neq (0, 0)$ be a period of tiling f . Without loss of generality we may assume that $b > 0$. Consider a horizontal strip of height b extracted from tiling f , e.g., the tiles $f(x, y)$ for $1 \leq y \leq b$. The sequences of horizontal colors on the top and the bottom of this strip are identical, with the horizontal offset a :

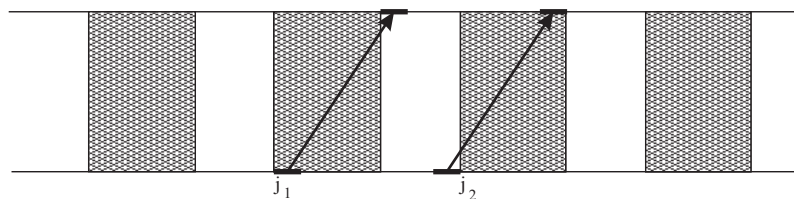


Within this strip, consider the rectangular $|a| \times b$ blocks

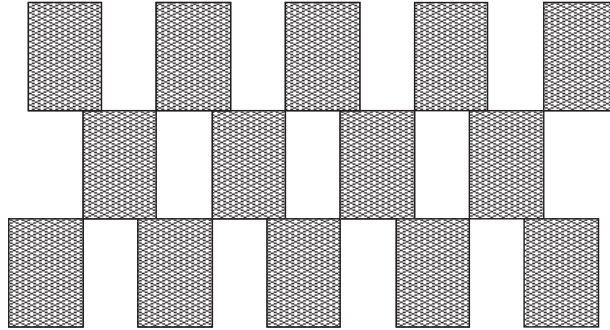


of tiles extracted from f with the bottom-left corner in position $(j, 1)$, for all $j \in \mathbb{Z}$. (And if $a = 0$ consider just the sequences of vertical colors on the b rows.) Since there are only a finite number of tiles in the protoset, there are only a finite number of such blocks. This means that for two different values of j , say j_1 and j_2 , the blocks are identical.

Now we can construct a valid periodic tiling of an infinite horizontal strip of height b by repeating the pattern between positions j_1 and j_2 . Note that the sequences of horizontal colors on the top and the bottom of this strip are again identical, with the horizontal offset a :



The tiling of the strip is valid and it has a horizontal period of length $j_2 - j_1$. A valid, two-way periodic tiling of the plane can now be obtained by stacking copies of the strip on top of each other, with the horizontal offset a :



□

4.2 Compactness principle

In later chapters we'll introduce a metric on tilings that induces a compact topology. This will imply several interesting results, but for the main algorithmic questions that follow next we just need to know that if a Wang set admits tilings of arbitrarily large squares then it admits a tiling of the whole infinite plane. This is a direct consequence of the compactness, but we state the result here without a direct reference to topology.

Let $\mathcal{P}$ be a finite set of Wang prototiles. Let us call any function

$$c : \mathbb{Z}^2 \rightarrow \mathcal{P}$$

a *configuration*, and let us denote by

$$\mathcal{P}^{\mathbb{Z}^2} = \{c : \mathbb{Z}^2 \rightarrow \mathcal{P}\}$$

the set of all configurations over the tile set $\mathcal{P}$. Note that configurations are arbitrary assignments of tiles on integer lattice points. i.e., the color constraints are not checked. Valid tilings are particular types of configurations.

Consider an infinite sequence $c_1, c_2, \dots$ of configurations, each $c_i \in \mathcal{P}^{\mathbb{Z}^2}$. We say that the sequence *converges* and $c \in \mathcal{P}^{\mathbb{Z}^2}$ is its *limit* if for every $(x, y) \in \mathbb{Z}^2$ there exists some $k \geq 0$ such that $c_i(x, y) = c(x, y)$ for all $i \geq k$. In other words: if we look at an arbitrary position and browse through a converging sequence $c_1, c_2, \dots$ then from some moment on we always see the same tile in that position. It is obvious that if a limit exists it is unique, and we denote this limit by

$$\lim_{i \rightarrow \infty} c_i.$$

A *subsequence* of $c_1, c_2, \dots$ is another sequence $c_{i_1}, c_{i_2}, \dots$ where $i_1 < i_2 < \dots$. A subsequence is hence obtained by picking infinitely many elements of the sequence, preserving their relative order. Obviously every subsequence of a converging sequence also converges and has the same limit.

The following theorem states the compactness of the configuration space:

Theorem 4.2 *Every sequence of configurations has a converging subsequence.*

Proof. Let $c_1, c_2, \dots$ be an arbitrary sequence, $c_i \in \mathcal{P}^{\mathbb{Z}^2}$. Let $\vec{r}_1, \vec{r}_2, \dots$ be some (arbitrary) enumeration of elements of $\mathbb{Z}^2$. In the following we show that there is a subsequence $c_{i_1}, c_{i_2}, \dots$ such that for every

$n \geq 1$, if $j \geq n$ then $c_{i_j}(\vec{r}_n) = c_{i_n}(\vec{r}_n)$, i.e., the subsequence has a constant value in the n 'th position $\vec{r}_n$ starting from the n 'th element of the subsequence. Then clearly the subsequence converges.

Let us choose indices $i_0 < i_1 < i_2 < i_3 < \dots$ inductively as follows: $i_0 = 0$ and $i_1 \geq 1$ is the smallest positive index such that there are infinitely many elements in $c_1, c_2, \dots$ that agree with c_{i_1} in the first position $\vec{r}_1$. Such c_{i_1} exists because there are only finitely many different tiles that can appear in position $\vec{r}_1$.

Suppose then that i_{k-1} has been chosen and we want to choose i_k for $k \geq 2$. We choose i_k to be the smallest integer that satisfies the following three conditions:

(A_k) $i_k > i_{k-1}$,

(B_k) $c_{i_k}(\vec{r}_j) = c_{i_{k-1}}(\vec{r}_j)$ for all $j = 1, 2, \dots, k-1$.

(C_k) There exist infinitely many indices i such that $c_i(\vec{r}_j) = c_{i_k}(\vec{r}_j)$ for all $j = 1, 2, \dots, k$.

Numbers i_k that satisfy (A_k)–(C_k) always exist for the following reasons: Because condition (C_k) was satisfied when i_{k-1} was chosen, we have infinitely many choices of i_k that satisfy (B_k). Set $\mathcal{P}^k$ is finite so there is a finite number of combinations of tiles that can appear in positions $\vec{r}_1, \dots, \vec{r}_k$. Consequently, among the infinitely many indices i_k that satisfy (B_k) there are infinitely many choices that also satisfy (C_k). Some of them hence satisfy all requirements (A_k)–(C_k).

It follows from properties (B_k) that $c_{i_1}, c_{i_2}, \dots$ converges: For an arbitrary $\vec{r}_n \in \mathbb{Z}^2$ all c_{i_j} for $j \geq n$ have the same tile in position $\vec{r}_n$. $\square$

Note: The proof is essentially the same as the proof of weak König's lemma which states that an infinite binary tree contains an infinite path. The proof did not require the axiom of choice. (The same result could also be easily proved using Tychonoff's theorem, but that is equivalent to the axiom of choice.)

Let us say that a configuration $c : \mathbb{Z}^2 \rightarrow \mathcal{P}$ tiles correctly at position $(x, y) \in \mathbb{Z}^2$ if $c(x, y)$ matches in color with its neighbors $c(x, y-1), c(x, y+1), c(x-1, y), c(x+1, y)$. A configuration is then a valid tiling iff it tiles correctly at each position.

The following corollary of the compactness principle states that if $\mathcal{P}$ can be used to properly tile arbitrarily large squares then it admits a valid tiling of the plane:

Corollary 4.3 *Let $\mathcal{P}$ be a finite set of Wang tiles. Suppose that for each finite set $F \subset \mathbb{Z}^2$ of positions there is a configuration that tiles correctly at each $(x, y) \in F$. Then $\mathcal{P}$ admits a valid tiling.*

Proof. Let $\vec{r}_1, \vec{r}_2, \dots$ be an enumeration of elements of $\mathbb{Z}^2$, and for each $n \geq 1$ denote

$$F_n = \{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}.$$

By the hypotheses of the corollary there exists for each n a configuration c_n that tiles correctly at positions $\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n$. By Theorem 4.2 the sequence $c_1, c_2, \dots$ has a converging subsequence. Let $c \in \mathcal{P}^{\mathbb{Z}^2}$ be its limit. Then c tiles correctly at every position $\vec{r}_k$ because there are arbitrarily large indices i such that c and c_i assign the same tile to position $\vec{r}_k$ and its neighbors. $\square$

4.3 Robinson's aperiodic tile set

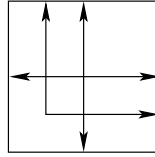
It is easy to construct Wang tiles that admit non-periodic tilings. For a long time it was thought that any finite set of prototiles that admits a non-periodic tiling must also admit a periodic one. This conjecture was refuted by R. Berger in 1966 when he constructed a set of Wang prototiles that only admit non-periodic tilings.

A finite set of prototiles is called aperiodic if

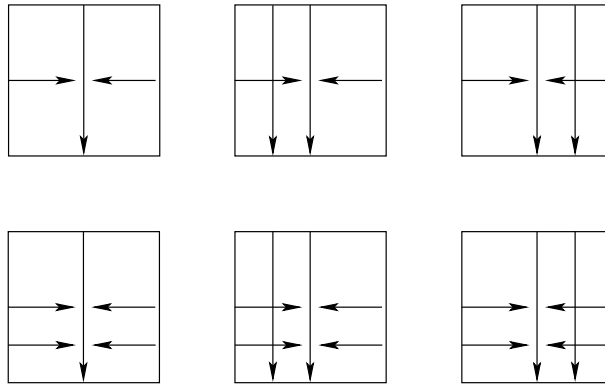
- (i) it admits valid tilings, and
- (ii) it does not admit any periodic valid tilings.

As an example of an aperiodic tile set we next describe a set of 56 Wang tiles due to R.M.Robinson. This set will be also useful later in our undecidability proofs. Instead of colors we use arrows to describe the matching rules between tiles. In valid tilings arrow heads and tails in neighboring tiles must match. This formalism can be easily converted into a color-based matching simply by assigning a different color for each orientation and positioning of arrows.

Robinson's tile set consists of tiles



called "crosses" and tiles



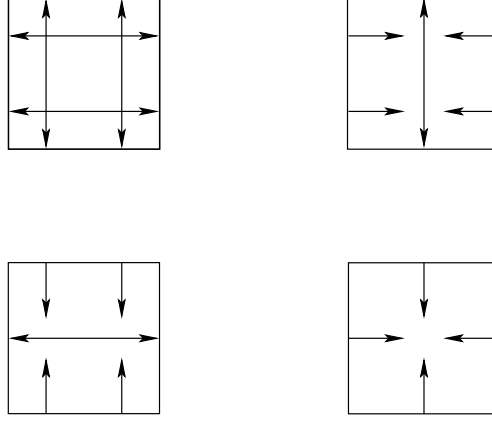
called arms. All tiles may be rotated so each tile comes in four orientations. Hence the total number of such tiles is 28.

The following terminology will be used:

- Every tile has central arrows at the centers of all four sides, and possibly some side arrows.
- A cross is said to face the directions of its side arrows.
- The arrow that runs through an arm is called the principal arrow of the arm, and the direction of the principal arrow is called the direction of the arm.

All six arms above are drawn in the north-to-south orientation. An important fact about arms is that if there are side arrows perpendicular to the principal arrow then these side arrows are towards the head of the principal arrow. Otherwise, all combinations of side arrows are allowed, as shown in the figure above.

We want to enforce a cross in the intersections of every other row and column. This can be established by forming the cartesian product ("sandwich tiles") with the parity tiles

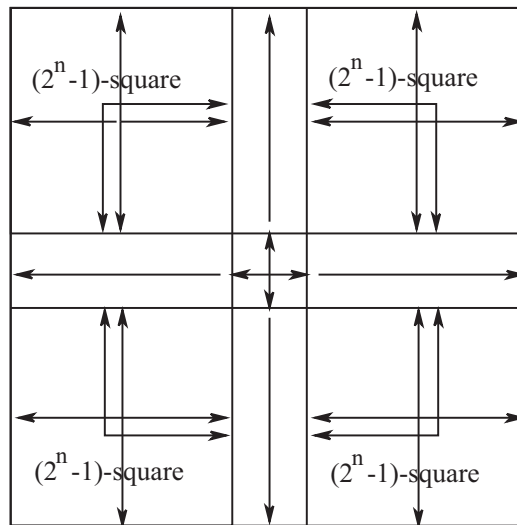


and by forbidding arms from the first parity tile. Since the only way the parity tiles tile the plane is by alternating the tiles on even and odd rows and columns, the first parity tile is forced at the intersections of every other row and column, and hence a cross is forced to appear in those locations. By numbering the rows and columns suitably we can assume from now on that all odd-odd positions of the plane contain a cross.

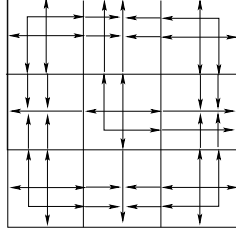
Note that between two crosses can only appear an arm, and the orientation of the arm has just two possible choices as it cannot point towards either cross. This means that the second parity tile only needs to be paired with north-to-south or south-to-north oriented arms, and the third parity tile is only paired with east-to-west or west-to-east oriented arms. The fourth parity tile is paired with any of the 28 tiles. So the final set contains $4 + 12 + 12 + 28 = 56$ different tiles.

Next we investigate valid tilings admitted by Robinson's tiles, and we show that the tile set is aperiodic. Specific patterns called 1-, 3-, 7-, 15-, $\dots$, $(2^n - 1)$ -squares are defined recursively as follows:

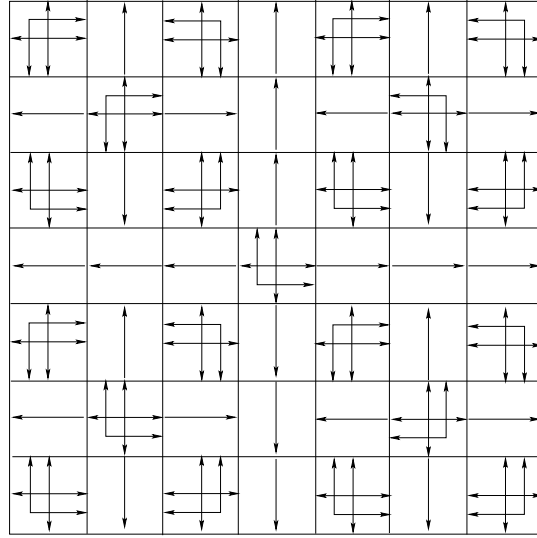
- (i) A 1-square is a cross at the odd-odd position,
- (ii) A $(2^{n+1} - 1)$ -square consists of a cross in the middle (in an even-even position), sequences of arms radiating out of the center and four copies of $(2^n - 1)$ -squares facing each other at the four quadrants:



Note that for every n there are actually four different $(2^n - 1)$ -squares as the cross at the center may be in any of the four possible orientations. For example, the following figure illustrates the 3-square facing north and east:

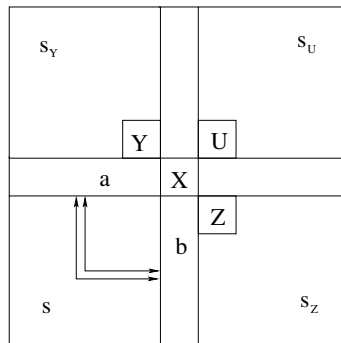


and the following figure shows the 7-square facing north and east. (For clarity, only the central, principal arrows of the arms are shown. The other arrows are uniquely determined by the orientations of the crosses.)



Inductively one easily gets the following properties of $(2^n - 1)$ -squares: (1) The tiling is valid within the square, (2) all edges on the border of the square have arrow heads pointing out of the square, so all edge neighbors of $(2^n - 1)$ -squares are forced to be arms, and (3) the only side arrows on the border are in the middle of the borders in the directions where the center cross of the square faces.

Consider an arbitrary valid tiling of the plane by Robinson's tiles. Let us show, using mathematical induction on n , that every cross in odd-odd position belongs to a unique $(2^n - 1)$ -square, for every $n = 1, 2, \dots$. The case $n = 1$ is trivial, as by definition 1-squares are themselves the crosses at odd-odd positions. Suppose then the claim is true for n and let C be an arbitrary cross in an odd-odd position. By the inductive hypothesis C belongs to a unique $(2^n - 1)$ -square s . There are four possibilities for the orientation of this square, but they are all symmetric. Let us assume without loss of generality that s faces north and east. In the following discussion we refer to symbols indicating positions in the following figure:



First we prove that tile X , outside the north-east corner of square s , must be a cross. Suppose the opposite: X is an arm. Then it has an incoming arrow on all but one side, so one of its edge neighbors in regions a or b must be an arm directed towards X . By continuing this reasoning we see that all tiles in one of the regions a or b must be arms directed towards X . But this means that the tile at the center of region a or b is an arm with an incoming side arrow at the wrong end of the principal arrow: the side arrows are only possible towards the head of the principal arrow. Hence the assumption that X is an arm must be incorrect, and X must be a cross.

Consider then tile Y that is a cornerwise neighbor of X . It is in an odd-odd position and therefore Y is a cross. According to the inductive hypothesis Y belongs to a $(2^n - 1)$ -square s_Y . This square cannot overlap with square s because then the tiles in the overlap region would belong to two different $(2^n - 1)$ -squares which contradicts the uniqueness property. Also the tile north of X cannot belong to s_Y because X is a cross. Hence Y has to be at the south-east corner of s_Y . Analogously, tiles Z and U are corners of disjoint $(2^n - 1)$ -squares s_Z and s_U , respectively. Tiles between these $(2^n - 1)$ -squares are forced to be arms radiating out from X . The side arrows at the middle of a and b force the center crosses of s_Y and s_Z to face squares s and s_U , so the squares of s, s_Y, s_Z, s_U and the tiles between them form a $(2^{n+1} - 1)$ -square that contains tile C .

We have proved the existence of a $(2^{n+1} - 1)$ -square that contains C . The uniqueness is obvious as the orientation of the (unique) $(2^n - 1)$ -square s that contains C determines the location of the center of the $(2^{n+1} - 1)$ -square that contains C .

We have proved that every 1-square belongs to a 3-square, which belongs to a 7-square, which belongs to a 15-square and so on. Based on this observation we can state:

Lemma 4.4 *Robinson's tiles form an aperiodic protoset.*

Proof. The $(2^n - 1)$ -squares are valid tilings of arbitrarily large squares, so a valid tiling of the plane exists (Corollary 4.3).

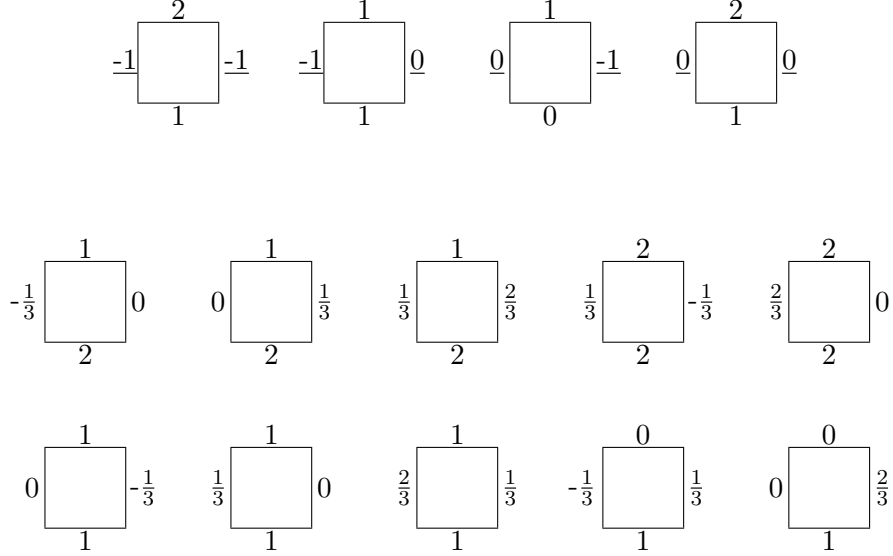
The centers of the quadrants of any $(2^n - 1)$ -square are crosses separated by $(2^{n-1} - 1)$ arms. As every valid tiling contains $(2^n - 1)$ -squares for every n , the tiling contains horizontally aligned crosses separated by arbitrarily long sequences of arms. So there can be no horizontal period, and a periodic tiling is not possible. □

4.4 An aperiodic set of 14 Wang tiles

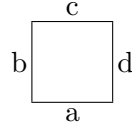
We have learned the aperiodic set of 56 Wang tiles by Robinson. In this section we learn a very different method of constructing aperiodic tile sets that yields a set with only 14 tiles, shown in the figure below. But note that even smaller aperiodic sets exist: E. Jeandel and M. Rao have an aperiodic Wang tile set that contains just 11 tiles, and they proved that 11 is the smallest possible size.

In our 14 tile set, the edges are labeled with rational numbers. Each number represents one color, so in valid tilings neighboring tiles must match in the numbers at the abutting edges. Notice also the the labels of the vertical edges of the first four tiles are underlined: This means that those numbers represent a different color than the same numbers without a line underneath.

The set consists of two parts: the first four tiles form the set $\mathcal{P}_2$ and the set of the last ten tiles is called $\mathcal{P}_{2/3}$. The aperiodic set $\mathcal{P}$ is the union of these two sets. As the vertical sides of the elements of the two parts have different labels, it is clear that on any valid tiling of the plane by $\mathcal{P}$, each horizontal row is tiled by tiles that come from $\mathcal{P}_2$ or $\mathcal{P}_{2/3}$ only.



The tiles perform arithmetic operations in the following sense: We say that tile



multiplies by q if $qa + b = c + d$. In other words, the tile multiplies the "input" number a on its bottom edge by q , adds the "carry forward" b from the left edge, and splits the result between the "output" c at the top edge and the "carry forward" d to the right. It is easy to verify that the tiles in $\mathcal{P}_2$ all multiply by 2, and the tiles in $\mathcal{P}_{2/3}$ multiply by $\frac{2}{3}$.

Consider a horizontal segment of n tiles that all multiply by the same number q . Let a_i, b_i, c_i and d_i be the numbers on the i 'th tile so that $qa_i + b_i = c_i + d_i$, for all $i = 1, 2, \dots, n$. Summing up over all n tiles we get

$$q \sum_{i=1}^n a_i + \sum_{i=1}^n b_i = \sum_{i=1}^n c_i + \sum_{i=1}^n d_i.$$

If the tiling constraint is satisfied then we have $d_i = b_{i+1}$ for all $i = 1, 2, \dots, n-1$, and if the segment also starts and ends with the same carry forward $d_n = b_1$ we have that

$$\sum_{i=1}^n d_i = \sum_{i=1}^n b_i.$$

This happens if the segment is extracted from a periodic tiling with horizontal period n . Then

$$q \sum_{i=1}^n a_i = \sum_{i=1}^n c_i.$$

Theorem 4.5 *The set $\mathcal{P}$ of 14 Wang prototiles above is aperiodic.*

Proof. We have two facts to prove: (i) no periodic tiling is possible, and (ii) some valid tiling exists.

(i) Suppose the opposite is true: there exists a periodic tiling $f : \mathbb{Z}^2 \rightarrow \mathcal{P}$. Then we know that such a periodic tiling must have a horizontal period h and a vertical period v , for some $h, v > 0$. (In fact we could choose these two numbers to be identical.)

Let $a_{i,j}, b_{i,j}, c_{i,j}$ and $d_{i,j}$ be the colors on the south, west, north and east edges of the tile $f(i, j)$ in position $(i, j) \in \mathbb{Z}^2$. It follows from the tiling rule that $d_{i,j} = b_{i+1,j}$ and that $c_{i,j} = a_{i,j+1}$. As discussed above, we have

$$q_j \sum_{i=1}^h a_{i,j} = \sum_{i=1}^h c_{i,j} = \sum_{i=1}^h a_{i,j+1},$$

where $q_j = 2$ or $\frac{2}{3}$ depending on whether the tiles on row j come from set $\mathcal{P}_2$ or $\mathcal{P}_{2/3}$. By combining these equations for rows $j = 1, 2, \dots, v$ we get the result that

$$q_1 q_2 q_3 \dots q_v \sum_{i=1}^h a_{i,1} = \sum_{i=1}^h a_{i,v+1} = \sum_{i=1}^h a_{i,1}.$$

It is clear from the tiles that we cannot have a horizontal row of tiles such that the bottom edges all have value 0, so we have $\sum_{i=1}^h a_{i,1} > 0$. Hence we can divide $\sum_{i=1}^h a_{i,1}$ from the equation, which leaves

$$q_1 q_2 q_3 \dots q_v = 1.$$

But each q_j is either 2 or $\frac{2}{3}$, and any product of these numbers is some power of 2 divided by a power of 3. Numbers 2 and 3 are relative primes, so no such product can equal 1, a contradiction.

(ii) It is enough to construct one valid tiling. The tiling will of course be non-periodic. We use the following notations and concepts. For any real number x , the floor $\lfloor x \rfloor$ of x is the largest integer not greater than x , that is, $\lfloor x \rfloor$ is the unique integer that satisfies $x - 1 < \lfloor x \rfloor \leq x$. Analogously, the ceiling $\lceil x \rceil$ is the smallest integer that is not smaller than x . By the *balanced representation* $B(x)$ of real number x we mean the bi-infinite sequence $\dots B(x)_{-1}, B(x)_0, B(x)_1, B(x)_2, \dots$ whose i 'th term is $B(x)_i = \lfloor ix \rfloor - \lfloor (i-1)x \rfloor$. Notice that the elements of the sequence are integers, and

$$\begin{aligned} B(x)_i &= \lfloor ix \rfloor - \lfloor (i-1)x \rfloor < ix - ((i-1)x - 1) = x + 1, \text{ and} \\ B(x)_i &= \lfloor ix \rfloor - \lfloor (i-1)x \rfloor > ix - 1 - (i-1)x = x - 1. \end{aligned}$$

Hence each element of the sequence $B(x)$ is either $\lfloor x \rfloor$ or $\lceil x \rceil$.

Consider an arbitrary real number $x \in [\frac{1}{2}, 1]$. We have that $0 \leq x \leq 1$ and $1 \leq 2x \leq 2$. The symbols in the balanced sequences for x and $2x$ are 0's and 1's, and 1's and 2's, respectively. Let us show that the prototiles of $\mathcal{P}_2$ admit a tiling of an bi-infinite horizontal strip whose bottom labels read the sequence $B(x)$ and the top labels read $B(2x)$. Let

$$\begin{aligned} a_i &= B(x)_i, \\ b_i &= 2\lfloor (i-1)x \rfloor - \lfloor (i-1)(2x) \rfloor, \\ c_i &= B(2x)_i, \text{ and} \\ d_i &= 2\lfloor ix \rfloor - \lfloor i(2x) \rfloor \end{aligned}$$

be the labels on the south, west, north and east edges of the tile in position $i \in \mathbb{Z}$ of the strip. It is clear from this definition that $b_i = d_{i-1}$ so the labels match on the tiling of the strip. Let us analyze values of a_i, b_i, c_i and d_i to prove that the tile with these labels is in our set $\mathcal{P}_2$.

From the properties of the balanced sequences we know that $a_i \in \{0, 1\}$ and $c_i \in \{1, 2\}$. Clearly,

$$\begin{aligned} d_i - b_i &= 2\lfloor ix \rfloor - \lfloor i(2x) \rfloor - (2\lfloor (i-1)x \rfloor - \lfloor (i-1)(2x) \rfloor) \\ &= 2(\lfloor ix \rfloor - \lfloor (i-1)x \rfloor) - (\lfloor i(2x) \rfloor - \lfloor (i-1)(2x) \rfloor) \\ &= 2a_i - c_i, \end{aligned}$$

so the tiles of the strip multiply by number 2. We also have

$$d_i = 2\lfloor ix \rfloor - \lfloor i(2x) \rfloor < 2ix - (2ix - 1) = 1,$$

and

$$d_i = 2\lfloor ix \rfloor - \lfloor i(2x) \rfloor > 2(ix - 1) - 2ix = -2.$$

Because d_i is an integer, the only possible values of d_i (and hence also b_i) are -1 and 0 . The following possibilities remain:

$$\begin{aligned} a_i = 0, c_i = 2 &\implies d_i - b_i = 2a_i - c_i = -2, && \text{not possible,} \\ a_i = 0, c_i = 1 &\implies d_i - b_i = 2a_i - c_i = -1 &\implies d_i = -1, b_i = 0, \\ a_i = 1, c_i = 2 &\implies d_i - b_i = 2a_i - c_i = 0 &\implies d_i = b_i = -1 \text{ or } d_i = b_i = 0, \\ a_i = 1, c_i = 1 &\implies d_i - b_i = 2a_i - c_i = 1 &\implies d_i = 0, b_i = -1. \end{aligned}$$

Only four possibilities exist, and these are precisely the four tiles in $\mathcal{P}_2$.

Next we analyze $\mathcal{P}_{2/3}$ in a similar way. Let $x \in [1, 2]$, so that $1 \leq x \leq 2$ and $\frac{2}{3} \leq \frac{2}{3}x \leq \frac{4}{3}$. The balanced representations of x and $\frac{2}{3}x$ consist of 1's and 2's, and 0's, 1's and 2's, respectively. Let us show that there is a tiling by $\mathcal{P}_{2/3}$ of a bi-infinite strip such that the labels on the bottom and the top of the strip read the balanced representations $B(x)$ of x and $B(\frac{2}{3}x)$ of $\frac{2}{3}x$. For brevity, let us denote $q = \frac{2}{3}$. The tile in position i of the strip has labels

$$\begin{aligned} a_i &= B(x)_i, \\ b_i &= q\lfloor (i-1)x \rfloor - \lfloor (i-1)(qx) \rfloor, \\ c_i &= B(qx)_i, \text{ and} \\ d_i &= q\lfloor ix \rfloor - \lfloor i(qx) \rfloor. \end{aligned}$$

The consecutive tiles of the strip match as $b_i = d_{i-1}$. We know that $a_i \in \{1, 2\}$ and $c_i \in \{0, 1, 2\}$. As above, we also have

$$\begin{aligned} d_i - b_i &= q\lfloor ix \rfloor - \lfloor i(qx) \rfloor - (q\lfloor (i-1)x \rfloor - \lfloor (i-1)(qx) \rfloor) \\ &= q(\lfloor ix \rfloor - \lfloor (i-1)x \rfloor) - (\lfloor i(qx) \rfloor - \lfloor (i-1)(qx) \rfloor) \\ &= qa_i - c_i. \end{aligned}$$

We also have

$$d_i = q\lfloor ix \rfloor - \lfloor i(qx) \rfloor < qix - (qix - 1) = 1,$$

and

$$d_i = q\lfloor ix \rfloor - \lfloor i(qx) \rfloor > q(ix - 1) - qix = -q.$$

Because d_i is an integer multiple of $\frac{1}{3}$, the only possible values of d_i (and hence also b_i) are $-\frac{1}{3}, 0, \frac{1}{3}$ and $\frac{2}{3}$. The following possibilities remain:

$$\begin{aligned} a_i = 1, c_i = 2 &\implies d_i - b_i = qa_i - c_i = -\frac{4}{3}, \\ &\text{not possible,} \\ a_i = 1, c_i = 1 &\implies d_i - b_i = qa_i - c_i = -\frac{1}{3} \\ &\implies d_i = -\frac{1}{3}, b_i = 0 \text{ or } d_i = 0, b_i = \frac{1}{3} \text{ or } d_i = \frac{1}{3}, b_i = \frac{2}{3}, \\ a_i = 1, c_i = 0 &\implies d_i - b_i = qa_i - c_i = \frac{2}{3} \\ &\implies d_i = \frac{1}{3}, b_i = -\frac{1}{3} \text{ or } d_i = \frac{2}{3}, b_i = 0, \\ a_i = 2, c_i = 2 &\implies d_i - b_i = qa_i - c_i = -\frac{2}{3} \\ &\implies d_i = -\frac{1}{3}, b_i = \frac{1}{3} \text{ or } d_i = 0, b_i = \frac{2}{3}, \\ a_i = 2, c_i = 1 &\implies d_i - b_i = qa_i - c_i = \frac{1}{3} \\ &\implies d_i = 0, b_i = -\frac{1}{3} \text{ or } d_i = \frac{1}{3}, b_i = 0 \text{ or } d_i = \frac{2}{3}, b_i = \frac{1}{3}, \\ a_i = 2, c_i = 0 &\implies d_i - b_i = qa_i - c_i = \frac{4}{3}, \\ &\text{not possible.} \end{aligned}$$

The ten possibilities are exactly the tiles of $\mathcal{P}_{2/3}$.

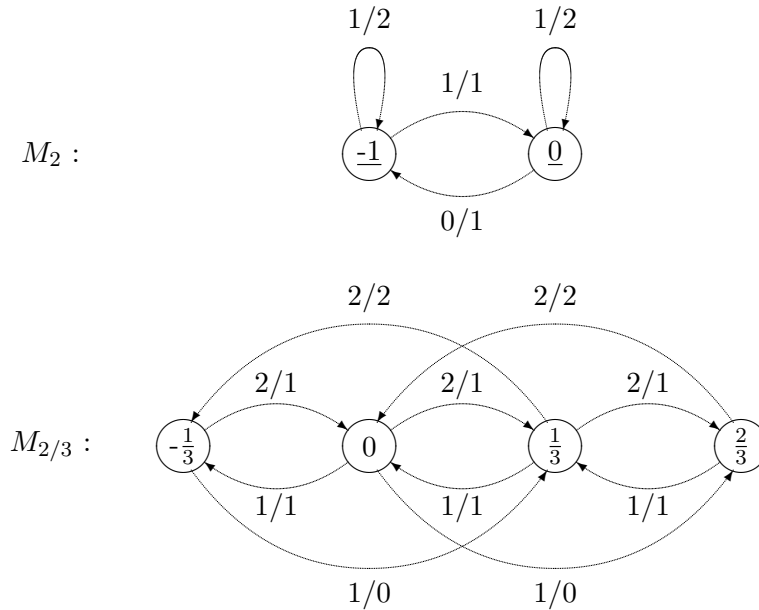
We have proved that for every $x \in [\frac{1}{2}, 1]$ we can tile a strip whose bottom and top edges read the sequences $B(x)$ and $B(2x)$, respectively, and for every $x \in [1, 2]$ we can tile a strip whose bottom and top edges read the sequences $B(x)$ and $B(\frac{2}{3}x)$. Define the following real function $f : (\frac{2}{3}, 2] \rightarrow (\frac{2}{3}, 2]$:

$$f(x) = \begin{cases} 2x, & \text{if } x \leq 1, \text{ and} \\ \frac{2}{3}x, & \text{if } x > 1. \end{cases}$$

It is easy to see that the range of f is the half-open interval $(\frac{2}{3}, 2]$, so function f is surjective. (In fact, function f is a bijection, but this fact is not relevant to the reasoning below.) It follows from the surjectivity of f that there exist bi-infinite sequences $\dots x_{-1}, x_0, x_1, x_2, \dots$ of real numbers such that $x_{j+1} = f(x_j)$ for all $j \in \mathbb{Z}$. In fact, since one element x_0 of the sequence can be chosen arbitrarily from the half-open interval $(\frac{2}{3}, 2]$, the number of such sequences is uncountably infinite.

As proved above, for each $j \in \mathbb{Z}$ we can tile an infinite strip whose edges read $B(x_j)$ and $B(x_{j+1})$. By stacking these strips on top of each other we obtain a tiling of the plane. In fact, we proved there exist uncountably many different valid tilings. □

The following diagram illustrates the tiles as a directed graph whose nodes are labeled by the vertical colors, and the edges are labeled by pairs of "input"/"output" symbols. Each edge corresponds to a tile: the tile with labels a , b , c and d is the edge from node b to node d that is labeled by a/c . Any bi-infinite path through the diagram that follows the edges gives a valid tiling of one bi-infinite strip. Such a diagram is called a finite state transducer.



Analogously, it is easy to construct for any given rational number q a finite set of tiles that multiply balanced sequences representing numbers of a given interval by q . The requested tiles have edge labels

$$\begin{aligned} a &= B(x)_i, \\ b &= q\lfloor(i-1)x\rfloor - \lfloor(i-1)(qx)\rfloor, \\ c &= B(qx)_i, \text{ and} \\ d &= q\lfloor ix\rfloor - \lfloor i(qx)\rfloor. \end{aligned}$$

where x is a number in the desired interval and $i \in \mathbb{Z}$. A simple analysis shows that we always have $-q < b < 1$ and $-q < d < 1$, so there only is a finite number of such tiles.

The balanced sequences we used in the proof have many interesting properties. Balanced sequences of rational numbers are periodic, but if x is irrational then $B(x)$ is non-periodic, but only "barely so": it is a so-called Sturmian sequence, which means that the number of different subsegments of length n is $n + 1$, for every n . This is the smallest possible number of different subsegments of length n in any non-periodic infinite sequence.

5 Undecidable problems concerning tiles

The following question (known as the domino problem or the tiling problem) arises naturally: How can one determine if a given finite set of Wang prototiles admits a tiling? Does there exist some simple (or even complicated) properties that one can use to develop a computer program to determine if a tiling is possible. The input to the program should be an arbitrary finite set of Wang tiles, and the output should be "yes" or "no" depending on whether the input admits a tiling. In this section we show that such a computer program does not exist. The non-existence is a mathematical fact that cannot be overcome by building more powerful computers or by developing new programming languages or tools. The undecidability will be deduced from Turing's result on the undecidability of the halting problem of Turing machines, using the reduction technique.

The tiling problem is an example of a decision problem. A decision problem is a problem that has an input parameter, and the answer to the problem is always "yes" or "no". When we fix the value of the input parameter we get an instance of the problem. An instance is called a "yes"-instance or a "no"-instance depending on whether the answer to the decision problem is "yes" or "no", respectively. For example, the problem "Does a given quadratic polynomial have a real root?" is a decision problem. Quadratic polynomials are instances (we always use the keyword "given" in the decision problem statement to indicate the input). For example, $x^2 - 2x + 2$ is a "no"-instance, while $x^2 + 2x - 1$ is a "yes"-instance to this problem. In the tiling problem, instances are finite sets of Wang tiles, and an instance is a "yes"-instance iff the protoset admits a tiling of the plane. The complement of a decision problem is the decision problem where the "yes" and the "no" answers have been switched: for example, the complement of the tiling question asks whether the given protoset does not admit a tiling. "Yes" means now that no tiling is possible.

An algorithm can be formally defined in various ways. In order to keep the discussion simple, we are going to leave it undefined. For our purposes it is sufficient to understand a (decision) algorithm to be a computer program that takes some input and returns a "yes" or a "no" answer on each input. We say that the algorithm solves a decision problem if the algorithm returns the correct yes/no-answer on every instance of the problem. If such an algorithm exists then the decision problem is called decidable and if no such algorithm exists then the problem is undecidable.

Strictly speaking, the input of a computer program is a string of bits (or a string over some other alphabet) so the instance of a decision problem has to be encoded into such a string. For example, Wang tiles could be encoded as sequences of four colors, each color represented as a unique binary string. There are of course many ways to do such an encoding, but all encodings are equivalent in the sense that the decidability status of the decision problem is not affected by the encoding. In our discussion encodings of inputs will be irrelevant as we are not going to write any actual programs – rather algorithms will be defined by describing in plain English the steps that the algorithm executes. The idea is that we all are sufficiently familiar with computer programming so that such a description (when detailed enough) will convince everyone that a computer program exists for solving the program. Notice also that when describing an algorithm we do not need to worry about things related to computing resources such as memory space etc. We are always supposed to have unlimited amounts of such resources.

For example, the following algorithm solves the question whether a given quadratic polynomial has a real root: Let $ax^2 + bx + c$ be the input to the algorithm. The algorithm starts by computing the discriminant $D = b^2 - 4ac$. Then it checks whether $D \geq 0$ or not. If $D \geq 0$ then the algorithm returns