APERIODIC TILINGS, ERGODIC THEORY, AND ROTATIONS

CHARLES RADIN
Department of Mathematics
University of Texas
Austin, TX 78712-1082, USA

Abstract We use ergodic theory to analyze the order and symmetry properties of certain structures, in particular bulk matter made of weakly interacting atoms, and polyhedral tilings of space. Statistical symmetries are emphasized as natural features related to the spectrum of such structures.

A Brief Overview

There is one underlying theme, the question: What kind of structures can you build from copies of a finite number of different components using local rules; more specifically, what qualitative order and symmetry features arise this way, (Radin, 1987; 1991)? (This includes, for example, classifying materials made from atoms, and classifying tilings made with matching rules.) Each section will focus on one key idea. In the first section it is the statistical mechanics approach to order and spectrum; in the second, it is the ergodic theory approach to substitution and finite type structures; and in the last section it is the role of rotations and statistical symmetry.

1. The Statistical Mechanics Approach to Order and Spectrum

We begin with a brief introduction to the physics of material structure.

It was suspected for centuries from studying features of the external shapes of crystals that such materials are made up of small constituents (called atoms) which are in special (highly ordered) configurations. When I say "such materials" I mean to be a bit restrictive, though much more general than what is normally thought of as "crystals"; basically, I refer to most of the inorganic solid matter in everyday surroundings.

R. V. Moody (ed.), The Mathematics of Long-Range Aperiodic Order, 499–519.

The model for understanding such macroscopic matter is called equilibrium statistical mechanics, and has been accepted for about a century. It has the following structure. Assume we wish to analyze the physical properties of a piece of such matter composed of many atoms (for simplicity we assume they are all of the same element), at various fixed temperatures $t > 0$ and chemical potentials c. (There is a slightly more complicated formula if we wanted to use a pressure variable, say, instead of chemical potential.) Assume there are N of these atoms contained in a cube L of volume $|L|$, which means their positions constitute N variable points in L, or a point in L^N.

We somehow assign a "(potential) energy" $E^L(x;c)$ to all possible configurations x of L^N; this will be discussed further below. (We are ignoring kinetic energy and quantum effects as their contributions to the following are relatively minor and easy to insert if desired.) Then for fixed $t > 0$ and c we consider

$$f_L(x) \equiv \frac{\exp[-E^L(x;c)/t]}{\int_{L^N} \exp[-E^L(x;c)/t]\, d^N x} \tag{1.1}$$

as a probability density for x. To emphasize macroscopic effects and minimize the effects of boundaries we take the (weak-$*$) limit of this probability measure as $L \to R^d$, getting a "Gibbs measure" $\mu_{t,c}$ on configurations in R^d. As we are interested in (solid) crystals, it is useful to take the weak-$*$ limit $t \downarrow 0$, giving $\mu_{0,c}$; the result of this is to eliminate thermal effects.

Let X be the collection of all countable subsets of R^d such that no two points are less than unit distance apart. Let H be the subcollection of finite subsets in X which contain the origin.

Define an "interaction" as a function $e : H \longrightarrow R$ such that $e(h) = 0$ if $|q| \geq r$ for some $q \in h$. (Here r is some fixed positive number, the "range of the many-body interaction," and it is reasonable to require that $e(h) = e(h')$ if h and h' only differ by a rigid motion.) For each cube L define E^L on X:

$$E^L(x;c) \equiv \sum_{q \in x \cap L} \left[\left\{ \sum_{h \subset T^{-q} x \cap B_r} e[h] \right\} - c \right], \tag{1.2}$$

where B_r is the open ball centered at the origin, of volume r, and T^q denotes translation by q. Finally, define $X_{\rm GS} \subset X$ by: $x \in X_{\rm GS}$ if for every L, and every $x' \in X$ such that $x \cap (R^d \backslash L) = x' \cap (R^d \backslash L)$, it follows that $E^L(x;c) \leq E^L(x';c)$. GS stands for "ground state," and the notation ignores the dependence on c.

Our objective is to show that for a reasonable interaction the configurations $x \in X_{\rm GS}$ consist of a crystal C in any possible position or orientation.

Now it can be proven (Cornfeld et al., 1982) under rather general conditions on E^L that $\mu_{0,c}(X_{\rm GS}) = 1$, so $\mu_{0,c}$ can be used, when convenient, to understand $X_{\rm GS}$. This is one of the main ideas of ergodic theory, and we will be taking advantage of it. Another point worth noting is that $X_{\rm GS}$ is by construction the solution set of an optimization problem; $x \in X_{\rm GS}$ if and only if x has, in the appropriate sense, the *lowest* possible energy (density).

It is not hard to check that $X_{\rm GS}$ is always nonempty, and closed under translation. We will postpone until the next section the definition of the natural topology for spaces like $X_{\rm GS}$; suffice it to say that two configurations in $X_{\rm GS}$ are "close" if they only differ in some large neighborhood of the origin by a small rigid motion. We note that $X_{\rm GS}$ is compact and metrizable in that topology (Radin and Wolff, 1992).

Consider the following example. Let $d = 2$, $c = 0$, and

$$e(h) \equiv \begin{cases} -25 + 24|h_1 - h_2|, & \text{if } h = \{h_1,\ h_2\} \text{ and} \\ & \qquad 1 \le |h_1 - h_2| \le 25/24; \\ 0, & \text{otherwise.} \end{cases} \tag{1.3}$$

If x is the triangular lattice as in Figure 1, then $x \in X_{\rm GS}$ (Radin, 1981). Also, if x' is the configuration in Figure 2, then still $x' \in X_{\rm GS}$. Notice the horizontal "fault line" in x'. The ergodic measures on $X_{\rm GS}$ each have as support the translations of x rotated by some fixed angle. (So $X_{\rm GS}$ may contain more than we want, but the rest is measure 0, and therefore negligible.)

Certain features of this example are important. We have proven (Radin, 1986), using variable parameters in the interaction, that generically among such models the set of ground states decomposes into closed, translation invariant "components"; the components are rotations of one another, as in the above "crystalline" example, and furthermore they are each *uniquely ergodic* with respect to translations—that is, there is one and only one translation invariant Borel probability measure on each of these components.

Let me summarize what we have. This model of matter consists of a space $X_{\rm GS}$ of states of the system, and a probability measure $\mu_{0,c}$ on it, for which the ergodic decomposition actually yields uniquely ergodic components, which are rotations of one another.

Physicists are "sure" that if an interaction were to be inserted in the above model, reasonable for some element, one should be able to "see" the appropriate crystalline structure C for that element in the measure $\mu_{0,c}$—although no such "reasonable" model has ever been successfully analyzed. Specifically, to "see" the crystal C, we look at many-body correlations, which are probability densities of events such as: "there is one atom at x_1, one at x_2, $\cdots$, and one at x_5". For such an event consider the positions

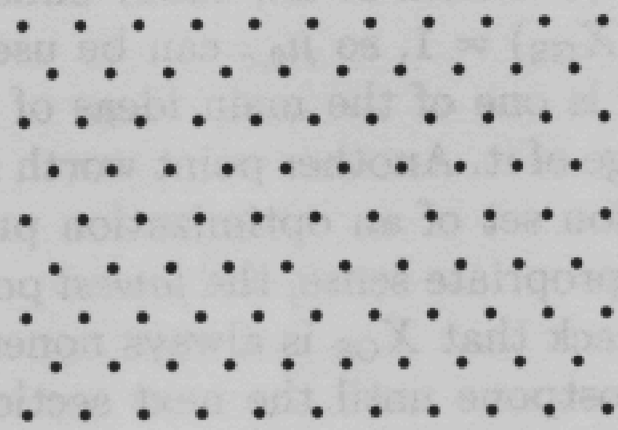


Figure 1. Ground state configuration

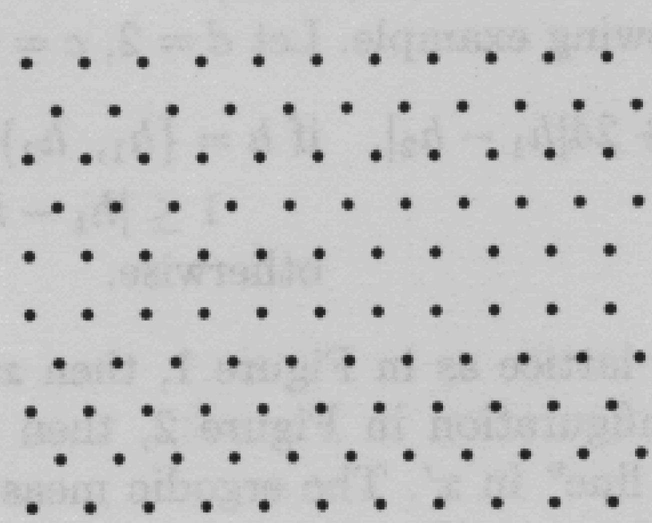


Figure 2. Configuration with fault line

$x_1 \cdots x_4$ fixed at sites of C, and leave x_5 variable. What we should find, assuming of course that the model indeed predicts the crystal C for the given interaction, is that the probability of this event is very small unless x_5 also is on a site of C, even very far from the others. The fact that the correlation senses the structure C even when the particles are far apart is called "long-range order" (Uhlenbeck, 1968).

This is the sense in which statistical mechanics predicts the crystal C; and it is not only the order properties of the crystal in which one is interested, but also the symmetries of C. (We note that for the most part physicists try to analyze such models at temperatures $t > 0$ not $t = 0$, because thermal effects are important for most phenomena.)

I have presented this statistical mechanical analysis as it could have been presented for many decades. Of course something changed in 1984 when quasicrystals were discovered. We still aren't sure what these beasts are, but many believe that they would fit simply in the above analysis as complicated examples of X_{GS}, determined by interactions appropriate for those alloys.

We now note a powerful technique widely used in quantum mechanics and ergodic theory. We have a topological space $X \equiv X_{\rm GS}$ on which the Euclidean group G is acting continuously, and a (probability) measure $\mu \equiv \mu_{0,c}$ on X which is invariant under the action of the group ($G \ni g : x \in X \to T^g x$). Therefore if we construct the complex Hilbert space $\mathcal{H}$ of complex functions on X square integrable with respect to μ, the natural induced map $G \ni g : f \in \mathcal{H} \to U^g f$, where $U^g f(x) = f(T^g x)$ defines a (strongly continuous) unitary representation of G. This allows us to use Stone's theorem (Riesz and Sz-Nagy, 1955) to get a convenient "spectral" representation of *some* of these unitary operators U^g, those corresponding to translations $g = q$:

$$U^q = \int_{R^d} \exp(-iq\lambda)\, dE_\lambda \tag{1.4}$$

or, taking "matrix elements,"

$$\langle f, U^q f \rangle = \int_{R^d} \exp(-iq\lambda)\, d\langle f, E_\lambda f \rangle, \tag{1.5}$$

where dE_λ (or $d\langle f, E_\lambda f\rangle$) is the spectral measure(s).

As a simple example consider the checkerboard tilings of the plane, which is uniquely ergodic if only translations of one tiling are in the space X. Then, with appropriate coordinates, $X = \{(\theta_1, \theta_2) \mid 0 \le \theta_k < 2\}$, and translation is $(T^q\theta)_k = \theta_k - q_k \mod (2)$. In this setting, the unique invariant Borel probability measure μ is normalized area. As before, $\mathcal{H} = L^2(X, \mu)$. Note that if we define, for all $j \in (Z/2)^2$,

$$\psi_j(\theta) \equiv \exp(2\pi i\, j \cdot \theta)$$

then $\psi_j \in \mathcal{H}$, $||\psi_j||^2 = 1$, and

$$U^q \psi_j(\theta) = \exp[-2\pi i\, j \cdot q] \psi_j(\theta).$$

So the ψ_j are normalized eigenvectors of the U^q. In fact, they form a (well known) orthonormal basis of $\mathcal{H}$. Therefore, for any $\psi = \mathcal{H}$,

$$\psi = \sum_{j \in (Z/2)^2} a_j \psi_j,$$

where $a_j = \langle \psi_j, \psi \rangle$. So

$$\begin{aligned} U^q \psi &= \sum_{j \in (Z/2)^2} a_j \exp(-2\pi i\, j \cdot q)\, \psi_j \\ &= \left[\sum_{j \in (Z/2)^2} \exp(-2\pi i\, j \cdot q) P_j \right] \psi, \end{aligned} \tag{1.6}$$

where P_j is the one dimensional projection $P_j\psi = \langle \psi_j, \psi \rangle \psi_j$. So equation (1.6) is Stone's formula equation (1.4), with dE_λ concentrated on $(Z/2)^2$.

This ***spectral representation*** is a very powerful tool. We will later use it to analyze rotational symmetry, but at this point we note two other uses for it, both related to order properties of the configurations in $X_{\rm GS}$. First of all, the concept of ***long-range order*** mentioned above is easily expressed in this language in terms of (in fact as the lack of) the "(strong) mixing" of the measure $\mu_{0,c}$, which is closely related to the condition that the spectral measures $d\langle f, E_\lambda f\rangle$ be absolutely continuous (Cornfeld et al., 1982).

The other point is the correspondence between the "X-ray diffraction patterns" off a configuration in $X_{\rm GS}$ and the measure $d\langle f, E_\lambda f\rangle$ for an appropriate f. Namely, if a configuration $x = \{q_j\} \in X_{\rm GS}$ represents the centers of atoms, all with the same electron density:

$$\rho_j(q) = k(q - q_j) = \begin{cases} 1/10 - |q - q_j|, & |q - q_j| \leq 1/10; \\ 0, & |q - q_j| > 1/10, \end{cases} \tag{1.7}$$

and $X_{\rm GS}$ is uniquely ergodic then defining $f : X_{\rm GS} \longrightarrow R$ by:

$$f(x) = \sum_{q_j \in x} k(-2q_j) \tag{1.8}$$

Steven Dworkin has shown (Dworkin, 1993) that if a plane wave of wavelength s in the direction W_0 is scattered off the atoms, the intensity I per unit volume scattered in the direction W is precisely $d\langle f, E_\lambda f\rangle$ with $\lambda = (W - W_0)/s$.

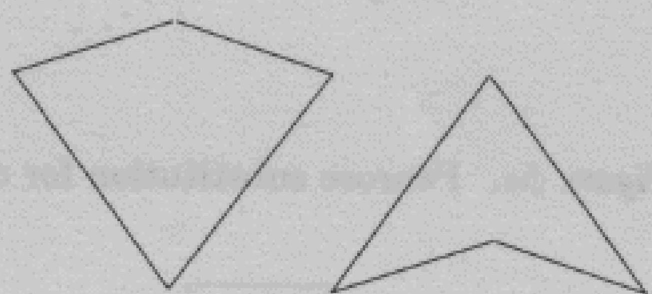
In conclusion, the concept of "order" and "diffraction spectrum" has been widely studied in physics and the statistical tools developed for it are normally expressed in terms of the spectral decomposition of the translation operators U^q, for the same reasons as in ergodic theory. We also emphasize that even though the object in this "crystal problem" is to understand the ubiquitous role of crystalline configurations, physical theory naturally leads to an embedding of that structure in an *ensemble* of structures with a (translation invariant) probability distribution on it, rather than a direct analysis of a single structure.

2. The Ergodic Theory Approach to Substitution and Finite Type Structures

Ergodic theory is concerned with a compact metric space X on which the group $G \equiv R^d$ is acting continuously ($G \ni g : x \in X \rightarrow T^g x$), and a (Borel) probability measure μ on X which is invariant under the action of the group. It is appropriate in this paper to think in terms of an example where X is a space of tilings, so I begin with such an example.

In tiling one starts with a finite "alphabet" $\mathcal{A}$, which is a collection of "letters," each of which is a polygon in the Euclidean plane. (There are obvious generalizations to higher dimensions, which will be used in the next section.) We define "tiles" as sets congruent to letters. Given an alphabet one defines the set $\mathcal{B}$ of "words," which are finite collections of tiles whose union is connected and simply connected, and in which tiles have pairwise disjoint interiors.

There may be many interesting ways to produce tilings; we will concentrate on two, which are called the methods of "substitution," and "finite type". We start with the substitution method, and as an example consider the usual alphabet of 2 letters in Figure 3 for Penrose's "kites and darts" substitution system (Gardner, 1977). (A part of a kite and dart tiling is shown in Figure 4.)



kite dart

Figure 3. Penrose letters for substitution

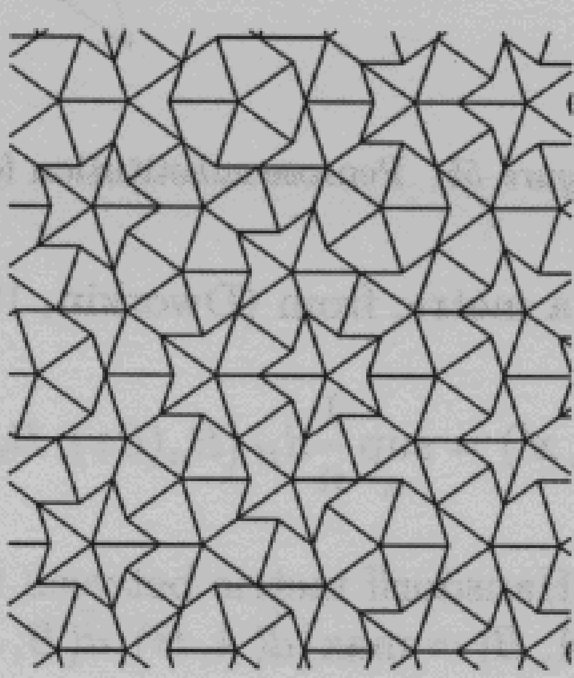


Figure 4. A Penrose tiling

Given the alphabet, we define the "substitution function" $F_P : \mathcal{A} \longrightarrow \mathcal{B}$ by Figure 5. This function is a composition of two maps; the first associates with each letter a collection of "shrunken tiles" (shown by solid lines in Figure 5), smaller than tiles by a factor $\lambda < 1$, in this case $\lambda = 2/(1+\sqrt{5})$, and the second map expands the shrunken word about the origin by $1/\lambda$. Although not immediately obvious, F_P extends naturally as a map from $\mathcal{B}$ to $\mathcal{B}$. With this notation we define the family $\mathcal{B}_{F_P}$ of words of the form $F_P^k(a)$, $a \in \mathcal{A}$, $k \geq 1$. Finally, we define the substitution system X_{F_P} associated with F_P to be the set of all tilings x of the plane such that every subword of x is congruent to a subword of some word in $\mathcal{B}_{F_P}$.

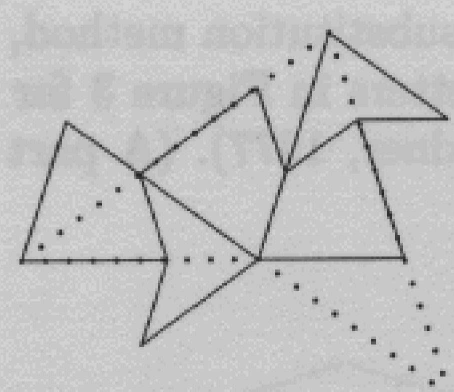


Figure 5a. Penrose substitution for dart

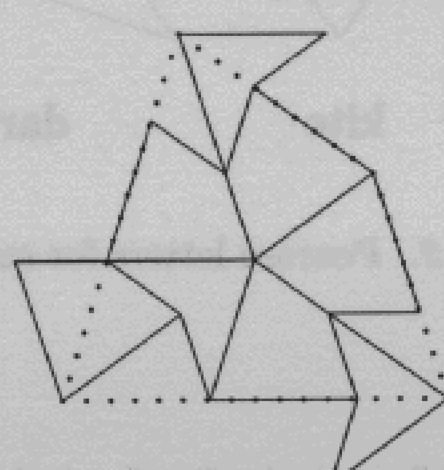


Figure 5b. Penrose substitution for kite

On X_{F_P} we define a metric from (Dworkin, 1994) as follows.

$$d(x, x') = \sup_n \frac{1}{n} d_H \left[B_n(\partial x), B_n(\partial x')\right], \tag{2.1}$$

where $d_H[A, B]$ is the Hausdorff metric between two compact subsets A, B of R^2, defined by $d_H[A, B] = \max\{\tilde{d}(A, B), \tilde{d}(B, A)\}$, where

$$\tilde{d}(A, B) = \sup_{a \in A} \inf_{b \in B} ||a - b||,$$

$||w||$ denotes the usual Euclidean norm of $w \in R^2$, and $B_n(\partial x)$ denotes the boundary of the largest subword of the tiling x which is a subset of the closed ball $\bar{B}_n \subset R^2$ of volume n centered at the origin. (It is easier to understand all this graphically than orally!)

For the ***pinwheel*** tilings the alphabet for the substitution system consists of a $1, 2, \sqrt{5}$ right triangle and its reflection, as in Figure 6. (A part

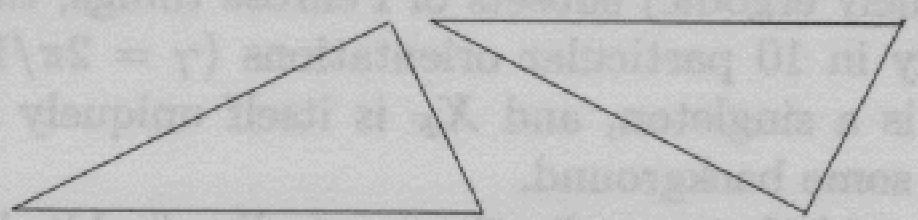


Figure 6. Pinwheel letters for substitution

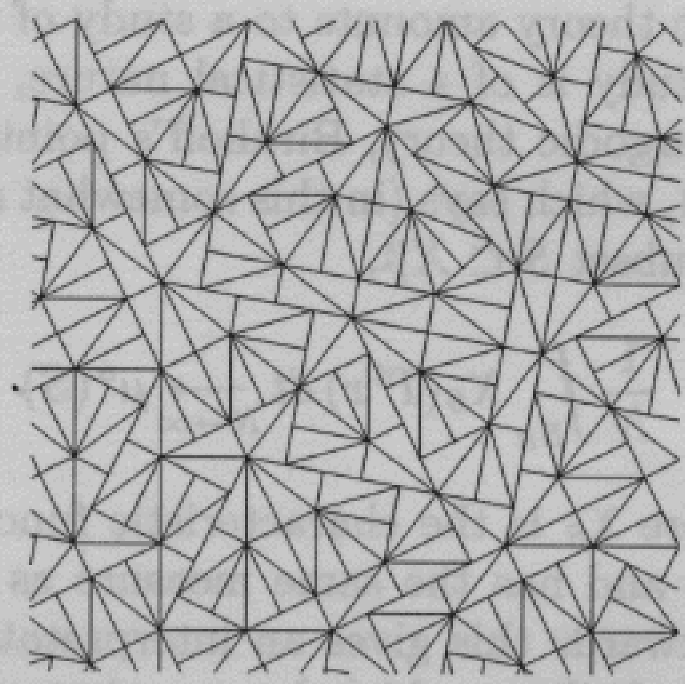


Figure 7. A pinwheel tiling

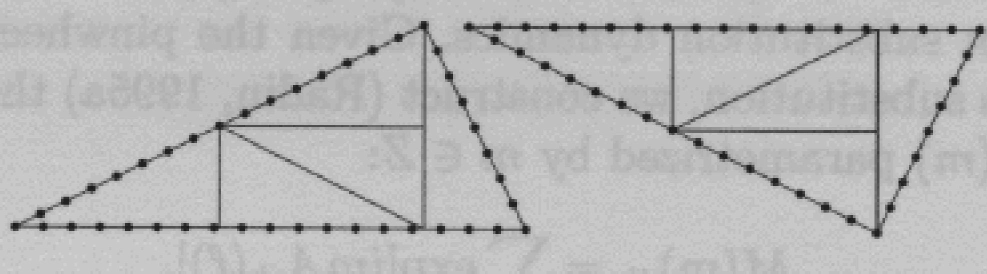


Figure 8. The pinwheel substitution

of a pinwheel tiling is shown in Figure 7.) The substitution function F_p, indicated in Figure 8, should again be considered a composite of two maps in the obvious way, and the substitution system X_{F_p} is defined as above. (In this case it is obvious that F_p extends naturally from the alphabet to the set of words).

Given a substitution system, under very general conditions X_F decomposes into closed, translation invariant families $\{X_F^\alpha : \alpha \in I\}$ of tilings which are uniquely ergodic; we will call the unique invariant measures μ^α. In general I is of the form $[0, \gamma]$ with addition modulo γ, and $R^\beta X_F^\alpha = X_F^{\alpha+\beta}$, where R^ω denotes the natural action on tilings of rotation about the origin by ω. (Note the similarity of this to the structure of ground states for generic interactions.) For instance, the space of Penrose tilings decomposes into nicer (uniquely ergodic) subsets of Penrose tilings, those in which the tiles appear only in 10 particular orientations ($\gamma = 2\pi/10$). For the pinwheel tilings I is a singleton, and X_F is itself uniquely ergodic. We will prove this after some background.

It is usually said that ergodic theory studies "orbits," that is families $\{T^t x : t \in R^2\}$. But when the "points" x themselves are configurations or tilings in the space in which you are translating, translation effectively just moves the origin around so that you are centered at different places in a "point". Then ergodic theory amounts to a study of the internal structure of the points. This study is of a statistical nature, depending as it does on the main tool of ergodic theory, Birkhoff's pointwise ergodic theorem (Cornfeld et al., 1982), which says (in this somewhat special situation) that for any nice enough subset $S \subset X_F^\alpha$,

$$\frac{1}{N}\int_{B_N} \chi_S(T^t x)\, dt \underset{N\to\infty}{\longrightarrow} \mu^\alpha(S) \tag{2.2}$$

for every $x \in S$, where χ_S is the characteristic function of S. (S is "nice enough" if it is open and has the same measure as its closure (Petersen, 1983).) Reading backwards this gives an interpretation of the measure of a set: if S is the set of tilings which have a given word at a given place in space (up to a small congruence), then $\mu^\alpha(S)$ is precisely the *frequency* with which that word appears in *any* tiling in X_F^α.

The proof that the pinwheel is uniquely ergodic extends a standard argument from substitution dynamics. Given the pinwheel alphabet of 2 letters, and its substitution, we construct (Radin, 1995a) the family of 2 by 2 matrices $M(m)$ parametrized by $m \in Z$:

$$M(m)_{jk} = \sum_\ell \exp[imA_{jk}(\ell)], \tag{2.3}$$

where $A_{jk}(\ell)$ is the angle (with respect to the standard letter) of the ℓ-th copy of the type j tiles in the substitution expansion of the type k letter: $1 \le j, k \le 2$. So with $s = \arctan(1/2)$, $M(m)$ is:

$$\begin{pmatrix} e^{-ims} + e^{-im(s+\pi)} & 2e^{-im(s+\pi)} + e^{-im(s+3\pi/2)} \\ 2e^{im(s+\pi)} + e^{im(s+3\pi/2)} & e^{ims} + e^{im(s+\pi)} \end{pmatrix}. \tag{2.4}$$

In light of the above analysis of the meaning of an invariant probability measure, we want to prove that every word has the same frequency density in all tilings and at all orientations. It is convenient to prove first that the density integrated over orientations is independent of the tiling. With some analysis one can show that it suffices to consider a word consisting of a single tile, and a region which is a high-level tile: $F_p^r(a)$ for arbitrary a and r. If a is type k, the number of type j tiles in $F_p^r(a)$ is $[M(0)^r]_{jk}$, where

$$M(0) = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}.$$

We now use the Perron–Frobenius theorem in the form (Ruelle, 1969):

Theorem 2.1 *Let M be a real $n \times n$ matrix such that:*

(a) $M_{jk} \geq 0$, for all j, k,

(b) $(M^N)_{jk} > 0$, for all j, k, for some $N \geq 1$.

Then, there is a simple eigenvalue $\lambda > 0$ for both M and its adjoint M^*, with corresponding eigenvectors ξ, $\tilde{\xi}$, with strictly positive components, and

$$\frac{M^p}{\lambda^p} x \xrightarrow[p\to\infty]{} \langle \tilde{\xi}, x \rangle \xi \quad \textit{for all } x \in C^n. \tag{2.5}$$

Corollary 2.1 *The spectral radius of M is λ.*

Corollary 2.2

$$\frac{(M^p)_{jk}}{\lambda^p} = \frac{\langle \hat{e}_j, M^p \hat{e}_k \rangle}{\lambda^p} \xrightarrow[p]{} \tilde{\xi}_k \xi_j > 0, \tag{2.6}$$

where $\{\hat{e}_j\}$ is the usual basis in C^n.

We use the Perron–Frobenius theorem for $M(0)$, obtaining

$$\frac{[M(0)^p]_{jk}}{\lambda^p} \xrightarrow[p]{} \tilde{\xi}_k \xi_j > 0.$$

So

$$\frac{[M(0)^p]_{jk}}{[M(0)^p]_{j'k}} \xrightarrow[p]{} \frac{\xi_j}{\xi_{j'}},$$

and this limit is nonzero for all k. This completes the proof that the relative population of the two tile types is well defined and independent of tiling.

Now by Weyl's criterion (Kuipers and Niederreiter, 1974), a sequence $\{x_j\}$ is uniformly distributed in $[0, 2\pi]$ if and only if for all $m \neq 0$

$$\frac{\sum_{k=1}^{K} e^{imx_k}}{K} \xrightarrow[K\to\infty]{} 0. \tag{2.7}$$

Noting how the orientations appear in $F_p^k(a)$, in order to prove that the orientations are uniformly distributed we need to show

$$\frac{[M(m)^p]_{jk}}{[M(0)^p]_{jk}} \xrightarrow[p]{} 0, \qquad \text{for all } m \neq 0. \tag{2.8}$$

We postpone the proof of this to the next section, when we also consider higher dimensional tilings.

We now consider a different method for producing tilings, the ***finite type*** method. Here one again starts with a finite alphabet $\mathcal{A}$ and the associated words $\mathcal{B}$. One then specifies a special set K of words, all of which have a tile with center of mass at the origin and are subsets of B_n for some fixed n. One then defines X_K as the set of *all* tilings x for which whenever you consider a ball of volume n centered at a tile the largest word in x inside that ball is congruent to a word in K. (It is usual to specify K by so-called "matching rules," or by thinking of the tiles as jigsaw puzzle pieces.)

For the kites and darts the finite type alphabet is Figure 9, and K_P is the restriction that the tiles fit together.

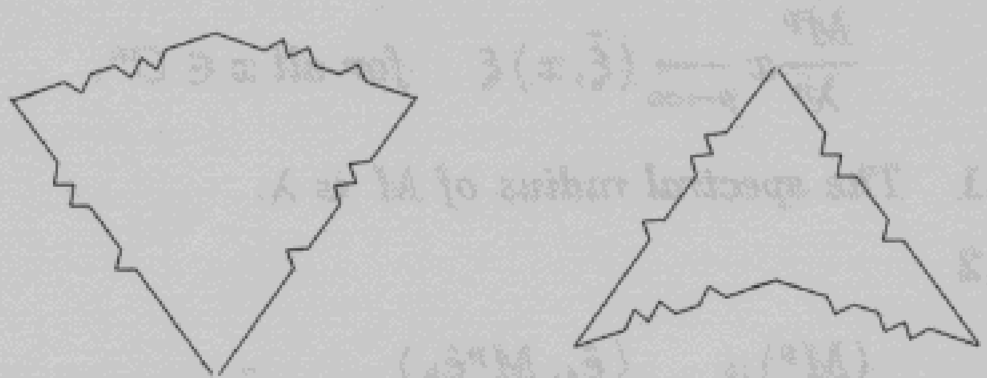


Figure 9. Penrose finite-type letters

It can be shown that the obvious identification between the letters of the substitution and finite type alphabets determines an isomorphism between X_{F_P} and X_{K_P} in any reasonable sense. For the ***pinwheel*** there is a large but finite alphabet $\mathcal{A}'$ such that after removal of a set of measure 0, X_{F_p} is isomorphic to X_{K_p}. (Again, K_p is the restriction that the tiles fit together.)

The point we wish to emphasize is that finite type systems can exist which are isomorphic to complicated substitution systems. This is surprising—it is certainly impossible to have a similar situation with 1 dimensional structures like sequences of symbols—and of course this was the point of (Wang, 1961) and (Berger, 1966) in creating the subject of aperiodic tilings. But what is even more amazing is the generality of this association. Mozes has given a *general procedure* (Mozes, 1989) which takes anything in a very wide class of substitution systems using colored squares, and gives an isomorphic finite type system (with much larger, finite alphabet)! That is, one starts out with a finite number of unit squares of different colors, and a substitution for them: for instance see Figure 10.

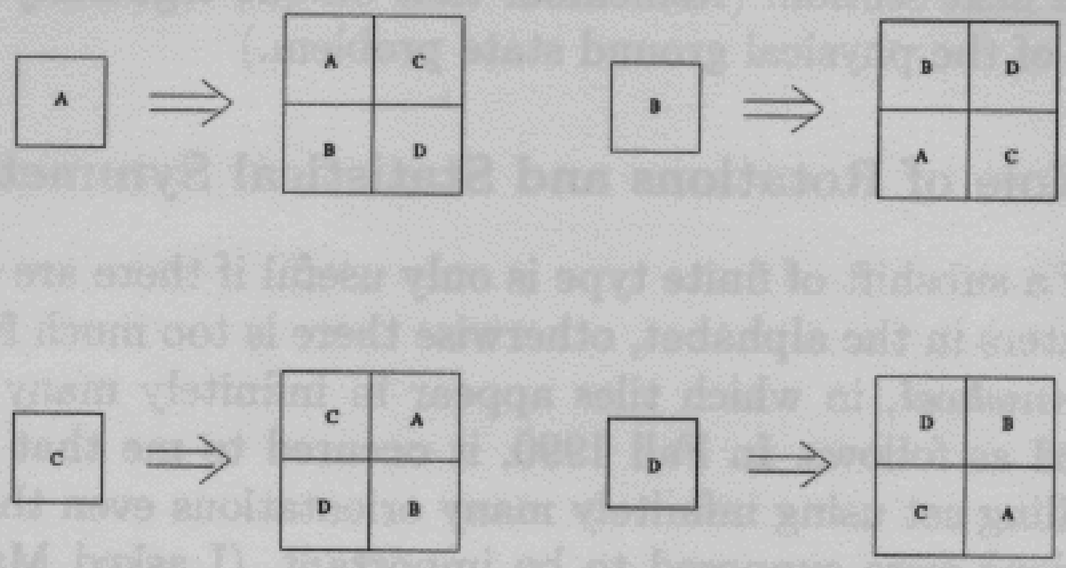


Figure 10. A substitution system

What Mozes does is modify the shapes (or colors) of the edges of the original letters, making many different "versions" of each of the originals, so that the tilings of the new finite type system are essentially in a natural one-to-one correspondence with the original tilings. Here "natural" means by ignoring the redundancy of the new letters, and "essentially" means ignoring a set of measure zero. And although the method of Mozes does not extend to shapes which are not basically square, the more complicated technique used for the pinwheel suggests that the correspondence is quite general indeed. (For special substitutions, such as the kite and dart, a finite type alphabet giving a tiling system isomorphic to the given substitution system is determinable by patches of some fixed size. However this is rarely possible—for example it is not possible for the pinwheel (Radin, 1996)—and the ***Theorem of Mozes***, and possible extensions, of course cannot rely on such a method.)

We conclude with the following crude observation. The method of analyzing structures that is ergodic theory depends on three things. First, you think of the structure you want to analyze as fixed in space so that for instance its translates are distinguished from it. A consequence of this is that translations can then be represented as actions on a space of structures (so you can use spectral decompositions etc.) Second you use a metric on the space of these structures such that two structures are close if they only differ a little in a large neighborhood of the origin. A consequence of this is that, effectively, translation just moves the origin around so you can concentrate on different portions of your structure. The last main element of ergodic theory is the pointwise ergodic theorem which gives you statistical information about the finite portions of your structure. It is particularly effective for uniquely ergodic systems for two reasons: there is no nonsense about "almost every" structure; and the convergence of equation (2.2) is uniform in the structure, not just pointwise. We will expand on this last point in the next section. (Remember that unique ergodicity is "justified" by analysis of the physical ground state problem.)

3. The Role of Rotations and Statistical Symmetry

The idea of a subshift of finite type is only useful if there are finitely many different letters in the alphabet, otherwise there is too much freedom. Consider the ***pinwheel***, in which tiles appear in infinitely many orientations. It originated as follows. In Fall 1990, it occured to me that I knew of no aperiodic tiling set using infinitely many orientations even though in some sense rotations were supposed to be important. (I asked Marjorie, as we all do, to find out what's known in this field, and she too didn't know any such example.) From the point of view of the crystal problem, there is certainly no reason to think this couldn't happen. But I even had trouble finding a finite number of polygons for which I could find *one* tiling with infinitely many orientations—let alone all their tilings. A colleague, Fillipo Cesi, found an example—a substitution system using 4 letters. With

$$x = \pi/7, \quad c = \cos(x), \quad s = \sin(x),$$

they are: two squares of side lengths 1 and $2 - c - s$; a rectangle with sides $c + s$ and $2 - c - s$; and a right triangle with legs c and s. The substitution is indicated in Figure 11.

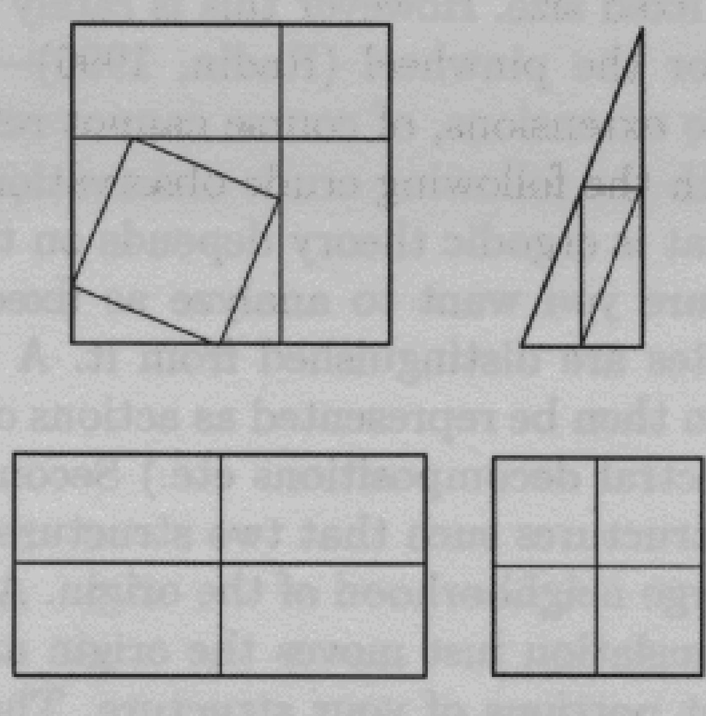


Figure 11. Cesi's substitution

For various reasons I was not very satisfied with Cesi's example; in particular, you must start the substitution with the large square to get anything interesting. The following Spring, John Conway was visiting Austin and I asked him if he knew a better way to accomplish this. In a few hours he came up with the pinwheel substitution (Figure 8). I knew about the work of Mozes and was pretty sure it was a general phenomenon, which meant there ought to be matching rules for this substitution. But because of the rotations I couldn't seem to adapt his method, and it took a couple of years to find (much more complicated) matching rules (Radin, 1994).

Now even though this model seems natural from the point of view of physics, one might wonder if in a way this amounts to allowing an infinite alphabet. Any "proof" one way or the other must lie in the results which follow this use of rotations, but so far it seems that this extra freedom does not in any way make it too easy to force complicated structures by local matching rules; in fact all it seems to do is import other parts of mathematics into tiling, for instance the algebra of the rotation group. That will be the point of this section.

Ergodic theory has been useful, using tools on ensembles of tilings rather than a single tiling. But there is one unusual aspect of the ergodic theory we are using, the unique ergodicity, which can be used to reverse the process, thereby using what we have learned about ensembles to apply to individual tilings. Fundamentally this is because in a uniquely ergodic system all tilings are essentially the same—that is, they have the same frequencies. This is immediate from the pointwise ergodic theorem as noted in the last section. We now exploit this fact.

Fix a tiling in a uniquely ergodic system, say a kite and dart (as a tiling in one of the uniquely ergodic components). To say that the frequencies are the same in all tilings in the system means that there really is no special origin; given any word, for instance a kite at a certain orientation, and $\epsilon > 0$, there is a size L such that in every region of size at least L the number of kites at that orientation divided by L is within ϵ of the "frequency" of that word.

Let's contrast this with the situation in an ergodic but not uniquely ergodic system, such as the "random" sequences of 0's and 1's: that is, consider a "typical" binary sequence, where typical means with respect to the usual product measure on sequences of 0's and 1's. For almost every sequence there are arbitrarily large blocks of 1's—so to estimate well the frequency of 1's around some origins you must take a much larger L than about other origins.

In this sense there really is less meaning for the origin in a kite and dart or pinwheel tiling than for random sequences. So let's see where we can get by ignoring the coordinate system in the plane. Fix a kite and dart tiling, in some (uniquely ergodic) component X of tilings. This property of the frequencies came about because of the *translation invariance* of the unique measure μ on X. Now let's look at rotations. If we rotate our tiling about some point by $2\pi/10$ we get another tiling in X. So the frequencies of any word (with fixed orientation) must be the same in the original and rotated tilings. That is, *the frequencies of finite words associated with any kite and dart tiling are invariant under rotation of the tiling by* $2\pi/10$. (This is simply another way of expressing the *rotational invariance* of the measure.)

The interesting thing about this property is that it is expressible without reference to other tilings, without reference to an ensemble of tilings. It is a (symmetry) property of this particular tiling. This is a basic new type of (rotational) symmetry of patterns in space, which we call ***statistical (rotational) symmetry*** (Radin, 1995b).

We can do the same for pinwheel tilings, but then we must use *frequency densities* rather than frequencies. That is, we consider the frequencies of a word in arbitrary but fixed *intervals* of orientations.

Let us take this a step further and consider a pinwheel tiling as a geometric space; assume you can move in the space only along edges of the triangles—that is, the space consists of the edges of these triangles. Lorenzo Sadun and I proved (Radin and Sadun, 1996) the following rather surprising fact: If you take any two points in this space which are a (Euclidean) distance N apart, then there is a path (along triangle edges) between the points of length $N + o(N)$. Another way of saying this is that the isoperimetric problem for this space has a circle for a solution (asymptotically).

This is not true of course for the usual checkerboard tilings, or for kite and dart tilings, or any other tilings I know. We found it surprising because for a "random" graph one would expect to need a path length of size $cN + o(N)$ for some $c > 1$, as for instance for a random direction on the checkerboard lattice, and the pinwheel is at least weakly mixing, so in that sense random.

This "roundness" property is potentially quite useful. Say you wanted to solve some problem numerically, for instance the Laplace equation in some elliptical region. The Laplace operator is locally symmetric under rotations—in fact this is an important property of this operator, giving rise to useful averaging properties of solutions of the Laplace equation. But using a square grid to do numerical simulations for this problem not only doesn't make use of this symmetry, it can in fact introduce false symmetries in the "solution". In other words, the vertices of a pinwheel tiling could be a useful grid to use for all sorts of planar models when the geometry is meaningful, in particular when (local) rotational symmetry is significant (Radin, 1995b).

There is a practical drawback to the above proposal. Since we are suggesting the use of statistical symmetry in finite portions of pinwheel tilings, let's analyze just how many different orientations we could expect in finite portions. For simplicity let's look at a finite portion which is a "letter of level k": that is, $F_p^k(a)$.

As noted in the last section, the orientations are obtainable from the matrix $M(m)^k$, where $M(m)$ is:

$$\begin{pmatrix} e^{-ims} + e^{-im(s+\pi)} & 2e^{-im(s+\pi)} + e^{-im(s+3\pi/2)} \\ 2e^{im(s+\pi)} + e^{im(s+3\pi/2)} & e^{ims} + e^{im(s+\pi)} \end{pmatrix} \tag{3.1}$$

The number of different orientations in a matrix element of this matrix is bounded above by the number of "words" of length k that can be made from the 6 different rotations in the matrix. If the rotations didn't commute there might be as many as 6^k different words available, but given the commutativity there can't even be as many as k^6. So the number of different orientations we would see, in the 5^k tiles that make up $F_p^k(a)$, is less than k^6; or in other words, it grows at most logarithmically in the number of tiles in the region.

This qualitative behavior cannot be improved for planar tilings made by substitution since it is due to the commutativity of the rotations in the plane. But how about higher dimensions?

John Conway and I have a nice 3 dimensional example (Conway and Radin, 1995). It is a substitution system built from a single tile, a prism that inflates to 8 copies of itself. We include in the inflation rotations by $2\pi/4$ and $2\pi/3$ about perpendicular axes. (These are not strange angles! In the pinwheel, the rotational symmetry was obtained by using an irrational angle, $\arctan(1/2)$. Here, it comes from the noncommutativity of rotations by simple angles.) We will get back to the question of the growth rate of orientations, but first we indicate how to analyze the uniform distribution of orientations.

We noted in the last section that for planar tilings such as the pinwheel Weyl's criterion suggests the construction of the family of matrices $M(m)$. Weyl's criterion generalizes to higher dimension (Kuipers and Niederreiter, 1974). In particular, a sequence of rotations $x_k \in \mathrm{SO}(n)$ is uniformly distributed if and only if

$$\frac{1}{K}\sum_{k=1}^{K} D^\lambda(x_k) \underset{K\to\infty}{\longrightarrow} 0. \tag{3.2}$$

for every nontrivial irreducible unitary representation D^λ of $\mathrm{SO}(n)$.

To use this it is convenient to go back a step. One can obtain the matrices $M(m)$ from the more basic substitution 2 by 2 matrix S with matrix elements in the group ring of SO(2) (with ring addition denoted $\oplus$):

$$\begin{pmatrix} (-s)\oplus(s+\pi) & (-s-\pi)\oplus(-s-\pi)\oplus(-s-3\pi/2) \\ (s+\pi)\oplus(s+\pi)\oplus(s+3\pi/2) & (s)\oplus(s+\pi) \end{pmatrix}, \tag{3.3}$$

where $s = \arctan(1/2)$. Namely, $M(m)$ is just the composition of an irreducible represention of SO(2) (the character $x \longrightarrow \exp(imx)$) with S. This is the route to take in general, giving a family of matrices S^λ labelled by the irreducible representations of SO(n).

To complete the proof for the pinwheel, we recall that we have to prove that

$$\frac{[M(m)^r]_{jk}}{[M(0)^r]_{jk}} \xrightarrow[r]{} 0, \qquad \text{for all } m \neq 0 \tag{3.4}$$

First we note that for all j,k and m, $|M(m)_{jk}| \leq |M(0)_{jk}|$. Assume that for some j,k, and some p, $|M(m)^p_{jk}| < |M(0)^p_{jk}|$ and that $M(0)$ is primitive (i.e., all coordinates are nonnegative and some power has all coordinates strictly positive), both of which are true for the pinwheel. We will show by contradiction that the spectral radius $\gamma(m)$ of $M(m)$ is less than that of $M(0)$, λ, for all m. Then $\gamma(m)^p$ is the spectral radius of $M(m)^p$ (Bachman and Narici, 1966), with corresponding eigenvector $\psi(m)$. Using the subscript + to signal replacement of the components of column vectors and matrices with their absolute values, we have

$$|\gamma(m)^p|\psi_+(m) \leq [M(m)^p]_+\psi_+(m) \leq [M(0)^p]\psi_+(m) \leq \lambda^p\psi_+(m) \tag{3.5}$$

Therefore $|\gamma(m)| \leq \lambda$. And if $|\gamma(m)| = \lambda$ then $\psi_+(m)$ must be the Perron–Frobenius eigenvector of $M(0)^p$ and $M(0)^p\psi_+(m) = [M(m)^p]_+\psi_+(m)$, so $[M(m)^p]_+ = M(0)^p$. This contradiction shows $\gamma(m) < \lambda$, and this proves our claim on the spectral radii of $M(m)$ and $M(0)$. From this, we note that

$$|M(m)^r_{jk}| \leq ||M(m)^r||, \qquad \text{for all } j,k,$$

and $||M(m)^r||^{1/r} \longrightarrow \gamma(m)$. So

$$\frac{|[M(m)^r]_{jk}|}{[M(0)^r]_{jk}} \leq \frac{||M(m)^r||}{[M(0)^r]_{jk}} \equiv a_r \tag{3.6}$$

We know $a_r^{1/r} \longrightarrow \gamma(m)/\lambda < 1$, so $a_r \longrightarrow 0$, as was to be proven.

For higher dimensional rotations this part of the argument would need to be improved. However we will only be concerned with a substitution which uses an alphabet of one letter, so we can ignore this complication.

Conway and I prove that the growth rate for words made out of the two rotations is algebraic instead of logarithmic. Technically, we prove that the group generated by these rotations is the free product of the two cyclic groups, amalgamated over a subgroup of finite index, and this gives us the growth rate. Actually we do a bit more. We define the group $G(m,n)$ generated by rotations about perpendicular axes by $2\pi/m$ and $2\pi/n$, $2 \le m \le n$. The groups corresponding to $m=2$ and $m=n=4$ are the only finite ones, and are easy to compute. We solve the next simplest cases: $G(3,3) = C_3 * C_3$ (the free product), and $G(3,4) = S_3 * (C_2)C_4$ (the amalgamated free product), where C_n is the cyclic group of order n and S_3 is the symmetric group of order 6. ($G(3,4)$ has presentation

$$\langle a, b, c : a^2, b^3, (ab)^2, c^4; a = c^2 \rangle,$$

the symmetric group S_3 has presentation

$$\langle a, b : a^2, b^3, (ab)^2 \rangle,$$

where $a = (12)$ and $b = (12)(23)$), and the cyclic group $C_4 = \langle c : c^4 \rangle$; in the amalgamation, $\{1, a\}$ and $\{1, c^2\}$ are identified.)

I find it amazing that $G(3,3)$ is a free product; and that this was not known! It is an indication of the richness of aperiodic tilings that it leads to such problems, quite naturally.

Let me summarize the points I hoped to get across in these three sections. The theme was: What kind of structures can you determine by local rules among finitely many different types of components?

Both statistical mechanics and ergodic theory have standard ways of approaching this problem, as ground states for finite range interactions and as subshifts of finite type. (Again we emphasize that these approaches embed the "single" configuration of interest in an *ensemble* of configurations, to use certain tools.) The existence of quasicrystals (1984) in the first case, and of aperiodic tilings (1966) in the second, raised unexpected possibilities.

It seems to be harder to create a traditional-style ("natural") interaction that has interesting ground states than to create an interesting aperiodic tiling set—at least since Wang and Berger led the way. So the main progress in our problem has come in aperiodic tilings, through a few examples and a theorem : Berger's founding example (1966) with "square" tiles; Penrose's kites and darts (1977) which started thoughts on symmetry and thus really transformed the problem from its original setting in logic to discrete geometry; the ***Mozes Theorem*** (1989) which suggested a general connection between substitutions and finite type systems, and confirmed the power of ergodic theory in the problem; and the ***pinwheel*** (1994), which led to ***statistical symmetry*** and other consequences of rotations.

The new 3 dimensional tiling (no matching rules yet!) may well start new directions also. But although new ideas seem to be coming at an increasing rate, I doubt the "crystal problem" will actually be "solved" anytime soon.

Acknowledgements

The author would like to acknowledge research support in part by the National Science Foundation Grant No. DMS-9304269 and Texas ARP Grant No. 003658-113.

References

Bachman, G. and Narici, L. (1966) *Functional Analysis*, Academic Press, New York.

Berger, R. (1966) The undecidability of the domino problem, *Mem. Amer. Math. Soc.*, **66**.

Conway, J. H. and Radin, C. (1995) Quaquaversal tilings and rotations, *Preprint*, Princeton University and University of Texas.

Cornfeld, I., Fomin, S., and Sinai, Ya. (1982) *Ergodic Theory*, Springer-Verlag, New York.

Dworkin, S. (1993) Spectral theory and x-ray diffraction, *J. Math. Phys.* **34**, 2965–2967.

Dworkin, S. (1994) The local structure and translational dynamics of tilings, *Preprint*, University of Texas.

Gardner, M. (1977) Extraordinary nonperiodic tiling that enriches the theory of tiles, *Sci. Amer.* (USA) **236**, 110–119.

Kuipers, L. and Niederreiter, H. (1974) *Uniform distribution of sequences*, Wiley, New York.

Mozes, S. (1989) Tilings, substitution systems and dynamical systems generated by them, *J. d'Analyse Math.* **53**, 139–186.

Petersen, K. (1983) *Ergodic theory*, Cambridge University Press, Cambridge, p. 271.

Radin, C. (1981) The ground state for soft disks, *J. Stat. Phys.* **26**, 365–373.

Radin, C. (1986) Correlations in classical ground states, *J. Stat. Phys.* **43**, 707–712.

Radin, C. (1987) Low temperature and the origin of crystalline symmetry, *Int. J. Mod. Phys. B* **1**, 1157–1191.

Radin, C. (1991) Global order from local sources, *Bull. Amer. Math. Soc.* **25**, 335–364.

Radin, C. (1994) The pinwheel tilings of the plane, *Annals of Math.* **139**, 661–702.

Radin, C. (1995a) Space tilings and substitutions, *Geometriae Dedicata* **55**, 257–264.

Radin, C. (1995b) Symmetry and tilings, *Notices Amer. Math. Soc.* **42**, 26–31.

Radin, C. (1996) Miles of Tiles, in *Ergodic Theory of Z^d-actions*, London Math. Soc. Lecture Note Ser. **228**, Cambridge University Press, 237–258.

Radin, C. and Sadun, L. (1996) The isoperimetric problem for pinwheel tilings, *Comm. Math. Phys.* **177**, 255–263.

Radin, C. and Wolff, M. (1992) Space tilings and local isomorphism, *Geometriae Dedicata* **42**, 355–360.

Riesz, F. and Sz-Nagy, B. (1955) *Functional Analysis*, (translated by L. F. Boron), Frederick Ungar, New York.

Ruelle, D. (1969) *Statistical Mechanics; Rigorous Results*, Benjamin, New York.

Sinai, Ya. (1982) *Theory of Phase Transitions: Rigorous Results*, Pergamon, Oxford.

Uhlenbeck, G. (1968) in *Fundamental Problems in Statistical Mechanics II*, (E. G. D. Cohen, ed.), Wiley, New York.

Wang, H. (1961) Proving theorems by pattern recognition II., *Bell Systs. Tech. J.* **40**, 1–41.