Acta Crystallographica Section A Foundations of Crystallography

ISSN 0108-7673

Received 1 September 2012 Accepted 25 February 2013

© 2013 International Union of Crystallography Printed in Singapore – all rights reserved

1. Certain shades of randomness

One tends to assume that point sets generated from random numbers do not exhibit particularly pronounced spatial patterns. However, as John von Neumann (1951) put it: *Any* one who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number. As a consequence, patterns emerge, where no one possibly would expect them.

This was shown, for instance, by Marsaglia (1968) regarding the sublattice structure of *pseudorandom* numbers generated by multiplicative congruential generators (MCGs), *i.e.* recurrence relations of the form $Z_{n+1} \equiv \mu Z_n \pmod{M}$ with integer multiplier μ and modulus M, $0 < \mu < M$ and $Z_i \in \mathbb{Z}/M\mathbb{Z} =$ $\{0, \ldots, M-1\}$. In comparison with Fig. 1(*a*), showing a perfectly ordered 8×8 square lattice, Fig. 1(*b*) depicts the



Figure 1

Two-dimensional point sets exhibiting distinct grades of randomness in their spatial distribution: (a) lattice ordered, (b) pseudorandom with pronounced sublattice structure, (c) quasirandom, (d) 'truly' random (*i.e.* pseudorandom with a sublattice structure beyond recognition).

Quasirandom guasicrystals

Institut für Materialphysik im Weltraum, Deutsches Zentrum für Luft- und Raumfahrt (DLR), 51170 Köln, Germany. Correspondence e-mail: wolfgang.hornfeck@web.de

Two-dimensional point sets derived from pairs of quasirandom numbers generated by the bit-reversal method introduced by van der Corput exhibit features well known from the quasiperiodic binary substitution tilings derived from the rhombic tilings of Penrose and Ammann–Beenker. The concept of geometric discrepancy, a measure describing the uniformity of distribution of quasirandom sequences or point sets, is discussed from the perspective of structural chemistry.

research papers

special (indeed, worst) case for a choice of multiplier and modulus as $\mu = 47$ and M = 65, respectively, yielding a similar sublattice of index 65 and exceptional high, *i.e.* square symmetry.

Although the pattern thus generated appears not random at all, *i.e.* it is just another square lattice inclined to the one depicted in Fig. 1(a), the distribution of points is still uniform. Successive numbers generated by the multiplicative congruential method appear as sufficiently random, at least in many applications, unless they are plotted in a way similar to the one shown in Fig. 1(b), where pairs of successive numbers divided by the modulus define the coordinates of a point in the plane.

Fig. 1(d), in contrast, depicts a truly *random* distribution of 64 points. Strictly speaking, 'random' here means pseudorandom, too, since the algorithm implemented in the postcript code generating the figure employs a pseudorandom number generator as well. However, in this case the generator is a more well behaving one, owing to special choices of its multiplier and its modulus. A common choice of parameters, describing the Park-Miller minimal standard MCG, may be given as $\mu = 7^5$ and $M = 2^{31} - 1$. In this case the Mersenne prime M_{31} is chosen as the modulus M, with the multiplier μ one of its primitive roots. This choice guarantees a most uniform distribution of lattice nodes regarding the inevitable sublattice structure of the MCG, therefore justifying their designation as pseudorandom. Moreover, the mere size of the modulus compared with the subset of 64 generated points conceals any visible hint to a particular sublattice structure.

The crystallographic implications of MCGs and their sublattice structure with respect to the coordinate description of crystal structures were explored by Hornfeck & Harbrecht (2009) and Hornfeck (2012).

Yet another *quasirandom* set of 64 points is shown in Fig. 1(c), which seems to share some of the regularity of a lattice arrangement as well as the irregularity of a random distribution. It is this point set *via* its corresponding planar tiling which exhibits remarkable features within a crystal-lographic context:

(i) The occurrence of pseudo-decagonal short-range order evoked by an unconventional combination of near-miss thin and thick rhombic Penrose tiles into a finite patch with a diamond tile at its centre and a slightly distorted decagonal convex hull.

(ii) The hierarchical arrangement of these pseudo-decagonal clusters into a fourfold, almost square, pattern, mediated by interwoven bands of near-miss thin and thick rhombic Ammann–Beenker tiles.

(iii) With the whole pattern generated by a most simple and general rule, namely a bit-reversal process, expressing a selfavoidance principle at its roots, thereby defining an infinite family of planar tilings.

Although the limiting cases of Figs. 1(a) and (d) seem well defined, one should be aware that, since the advent of quasicrystals, a consistent concept of long-range order is missing. Accordingly, quite distinct notions of randomness exist. Notably a truly random distribution of points shows a pronounced *clustering* of points accompanied by the corresponding less densely occupied regions (voids). Often regarded as counter-intuitive, this is the logical outcome of the conditions that the placement of each point is independent of the placement of the ones before and that each position shares an equal probability to be occupied by a point. Truly random numbers are distinguished by sharing the properties of being unpredictable, uncorrelated and unbiased (Hayes, 2011). For instance, a single throw of a fair dice yields any of the possible outcomes from one to six with an equal probability of 1/6, each throw is independent from those before, and in the limit of a large number of repeats the frequency of occurrence of each event will be uniformly distributed yielding the expectation value of (1 + 2 + 3 + 4 + 5 + 6)/6 = 3.5.

Attenuating the conditions on predictability, independence and equidistribution yields numbers that still appear as random, with respect to certain statistical tests in favour, but nevertheless are truly deterministic. The conceptual difference between pseudo- and quasi-random numbers relative to one another and with respect to genuine random numbers originates from the exact way and extent to which the conditions for randomness are diminished and bearing in mind the kind of application for which the random numbers should be used for. For instance, any inhomogeneity regarding the distribution of points appears especially disadvantageous in certain applications of random numbers, such as Monte Carlo methods which rely on the randomized but effective sampling of some possibly high-dimensional parameter space [see Hammersley & Handscomb (1975) for a general overview and Hayes (2011) for a discussion of practical examples]. Thus, the uniform distribution of the randomly generated sampling points is a property of paramount importance [for a survey on the uniform distribution of sequences see Kuipers & Niederreiter (1974)]. Random numbers fulfilling this criterion are denoted as quasirandom, to distinguish them from their pseudorandom relatives.

Similar sampling problems are found in crystallography, *e.g.* concerning probabilistic approaches to the solution of the phase problem in structure determination. An extended

discussion on this topic may be found in the Appendix B to this article.

Since the literature on random number generation is extensive we restrict ourselves to reference the monographs of Niederreiter (1992) and Hellekalek & Larcher (1998), both of which focus on pseudo- and quasirandom numbers, lowdiscrepancy sequences and point sets, their construction and application for Diophantine approximation, numerical integration and Monte Carlo methods in general.

2. Patterns in the quasirandom realm

In the following, we aim to give an exposition of a special case and its relation to an infinite family of planar tilings (\S 2). On this occasion, we summon some well established concepts from the field of discrete mathematics which may be of some use for the structural chemist or crystallographer (\S 3).

2.1. Pseudo-decagonal short-range order from bit reversal

Visual inspection of the pattern constructed from a set of *quasirandom* points and depicted in Fig. 1(c) exhibits features characteristic of *quasicrystals*, in particular point configurations of pseudo-*decagonal* local symmetry. A single pseudo-decagonal cluster is highlighted in Fig. 2(a), where the points are now identified with the vertices of a planar tiling solely consisting of quadrilaterals, *i.e.* slightly varying thin and thick rhomb(oid)s. Four adjacent quadrilaterals always share a common vertex (except for borderline tiles). Another three decagonal clusters are situated in the remaining quadrants.

The point set under consideration is only one example out of a larger class of sets, known as Hammersley sets, which are based on and generalize ideas introduced by van der Corput (1935) regarding the construction of certain number sequences (low-discrepancy sequences) now bearing his name. The associated planar tilings therefore may also be



Figure 2

Hammersley point sets of size M = 64 and M = 256 with an underlying near-miss binary rhombic tiling and pseudo-decagonal clusters in an almost square arrangement (*a*), local distortions in between decagonal subsets (*b*), and a series of vertex figures consisting of a quadruple of alternating pairs of squares and rhombi (*c*).

termed *Hammersley* tilings and, for those special cases exhibiting pseudo-decagonal features, designated by their own symbol H_n .

Details of the construction scheme devised by van der Corput are given in Appendix A1. It follows from the construction principle that a van der Corput sequence may be calculated up to any desired length $M = N \in \mathbb{N}$. However, Hammersley tilings H_n do occur only for certain values of M. Since the bit-reversal process for the point sets under consideration is rooted in base two, a natural selection scheme exists based on their size, in which M is given as a power of two $M = 2^{m}$. Empirically one finds pseudo-decagonal clusters only for cases where *m* is even, yielding the restriction $M = 2^{2n}$. In order to designate the first instance, for which a pseudodecagonal cluster appears, with an index n of unity, the formula is changed to $M = 2^{2n+2}$ or $M = (2^{n+1})^2$ in order to emphasize that the number of vertices is likewise a power of two and a square number. These numerical restrictions relate to the existence of a bit-reversal permutation for each Hammersley tiling H_n , which may be used to encode the spatial information of its vertex set and prove the generality of certain of its features (cf. Appendix A3).

Fig. 2(*b*) shows a larger set of 256 points exhibiting similar features on a different scale, emphasizing the general nature of the previous observation. Again, a single decagonal motif out of a total of 16 is highlighted and, in addition, the nine interstitial regions separating the decagonal clusters are emphasized. These 'gaps' at the barycentres of four adjacent decagons denote the location of the rhomb(oid)s exhibiting the largest and most variable distortions, in terms of the extremal values their acute and obtuse angles adopt. In a crystallographic context this local variance of a structural motif within the surroundings of an ordered framework resembles an (in)commensurate *modulation* and raises the question regarding the possibility to describe the whole pattern in this way, using an average unit-cell approach.

This effect proceeds for points sets of progressively larger size, as can be inferred in continuation of the discussed examples for $M = 2^4$ (cf. Table 1), $M = 2^6$ (cf. Fig. 2a) and $M = 2^8$ (cf. Fig. 2b), especially taking into account the general construction principle (Appendices A1 and A2) and its iterative nature (Appendices A3 and A4).

2.2. Tilings composed of regular squares and rhombs

Upon augmenting the patch of Fig. 2(b) into an infinitely extended tiling of the plane, every vertex has *four* incident edges with an overall similar vertex configuration of alternating thin and thick rhomb(oid)s [here the notation 'rhomb(oid)' reflects the distinction between ideal and less ideal geometric shapes: a rhomb defined by a single edge length as opposed to its distorted form, a rhomboid, a parallelogram with two distinct edge lengths].

Thus, the tiling is topologically equivalent to a regular tiling of the plane by squares ($4^4 = Q^4$; Fig. 2*c* left) and to square tilings derived from the Fibonacci and Octonacci sequences although the latter also contain isosceles trapezoidal tiles [see Lifshitz (2002) for the Fibonacci tiling and Sire *et al.* (1989) for the Octonacci one]. Fig. 2(c) depicts a sequence of vertex figures, each consisting of a quadruple of alternating thin and thick rhomb(oid)s, which all admit periodic tilings of the plane, since opposite pairs of concave and convex edge paths match.

Fig. 2(c) thereby facilitates a visual comparison of the relative magnitude of the angles encountered in the rhombic tiling compared with more ideal vertex configurations, starting from the limiting case of the regular square tiling (a square denoted by Q with 'acute' angle $\alpha = 90^{\circ}$) and proceeding with the special cases in which a pair of squares is alternatingly combined with congruent pairs of either thick Penrose rhombs (R, acute angle $\alpha = 72^{\circ}$), equilateral triangles (T, 60°), Ammann–Beenker lozenges (ℓ , 45°) or thin Penrose rhombs ($r, 36^{\circ}$). The tiling may be derived as well from a semi-regular Archimedean tiling constituted of equilateral triangles and squares (snub square tiling; vertex figure $3^2434 = T^2QTQ$), which can be continuously transformed into a tiling of rhombi (*i.e.* angular distorted squares) and squares by disregarding all edges shared by adjacent equilateral triangles.

A case study into the diffraction properties of tilings of this kind was recently reported by Baake & Grimm (2012) relating them to what they call 'planar σ -phases', *i.e.* a mathematical abstraction resembling a special type of Frank–Kasper phase, which represents a well known example from the structural chemistry of intermetallic compounds sharing certain features of quasicrystals! Moreover, by means of adjusting the angular distortion to special values such as $\phi = \pi/8$ or $\phi = \pi/12$ they were able to calculate diffraction patterns with pseudo-octagonal and pseudo-dodecagonal intensity distributions (*cf.* §3.3 and Figs. 6 and 7 of their work)!

Finally, uniform equitransitive tilings exist where the rhombi may be dissected into a pair of isosceles rather than equilateral triangles (see Grünbaum & Shephard, 1987; p. 88, Fig. 2.5.7b), thereby bearing a closer resemblance to the tiling analysed in this work.

2.3. Relations to the tilings of Penrose and Ammann-Beenker

A more quantitative relation to the rhombic tilings of Penrose and Ammann–Beenker is established by an analysis of the geometric properties of the constituent tiles. For this purpose the tiling is decomposed according to the distinguishable classes of equivalent tiles, resulting in the colouring shown in Fig. 3.

The finite tiling derived from the set with 64 quasirandom points consists of a total of six distinct types of tiles, up to enantiomorphism, whose geometric properties are listed in Table 1. Of these, three tiles are rhombs, thereby defining a total of *three* distinct distances, while the other three are rhomboids.

Three tiles may be designated as *thin* rhomb(oid)s, regarding the criterion that their acute angles are less than 60° . Since one tile approximates a rhombus consisting of a pair of equilateral triangles, the remaining two tiles are classified as *thick* rhomb(oid)s accordingly.



Figure 3

Finite patch of the tiling introduced in Fig. 2 with a colouring corresponding to the six distinct types of tiles (up to mirror symmetry according to point group 2m) as well as the subtle deviation from a square arrangement. A wedge between the origin and its first successing point to the right is used to indicate the slight offset of repeating motifs. The red, magenta and green tiles are rhombs, in which case all edges are of equal length, whereas for the yellow, cyan and blue tiles two distinct edge lengths occur, as would be the case for rhomboids.

An even finer classification, taking into account the actual values for the acute and obtuse angles, establishes the connection to quasiperiodic tilings, since both thin and thick Penrose rhombs as well as Ammann-Beenker type lozenges and squares may be assigned to their slightly distorted counterparts in the actual tiling. For instance, the thin and thick rhombs highlighted in grey in the upper left part of Fig. 2 are only slightly deformed variants of their counterparts in the rhombic Penrose tiling with acute and obtuse angles of 36.9° (36°) and 143.1° (144°) for the *thin* rhomb and 77.5° (72°) and 102.5° (108°) for the *thick* one, respectively, with their ideal values given in brackets. The corresponding values for the Ammann-Beenker tiles are given in Table 1. Overall, the maximum geometrical deviations from an averaged edge length of $\sqrt{79}/64$ as well as the ones with respect to the ideal angles (given in brackets) are smaller than 8%, and even smaller than 1% in the case of single tiles. In general, the thin rhombs yield a considerably better match compared with their ideal counterparts than the thick ones.

The six distinct types of tiles arrange themselves into a pattern of 49 tiles in total, with seven tiles in each row and column. Thin and thick Penrose-like rhomb(oid)s are present in equal amounts ($16 \times$ each), whereas twice as many thick Ammann-Beenker-like rhomboids occur as thin ones $(8 \times$ *versus* $4 \times$). Diamond-like rhombs, forming the centre tiles of a decagonal cluster, occur four times. A single tile located at the centre of the tiling remains unaccounted for. However, in some way this seems to be a logical consequence, since neither of the other tiles follows their conventional matching rules valid within a quasiperiodic tiling, thus demanding the occurrence of gaps in the tiling. Furthermore, the same rhomb is constituted of two points which take no part in the formation of the pseudo-decagonal clusters (they are also fixpoints of the corresponding permutation, cf. Appendix A3, but this does not hold in general). From a structural chemist's perspective such entities appear as 'glue atoms'. This analogy

Table 1

Edge lengths $|e_i|$ and angles α_i for the set of six distinct coloured tiles as illustrated in Fig. 3.

The type of tile is indicated by letters: R/r = Penrose thick/thin rhomb, Q/ℓ = Ammann–Beenker thick/thin rhomb, T = equilateral triangle, c = central thin rhomb.

Colour	$ e_1 /64^{-1}$		$ e_2 /64^{-1}$	α(°)	α	Type
	1011/01		1021/01	acute ()	Cobtuse ()	1990
Red	$\sqrt{80}$	=	$\sqrt{80}$	36.9 (36)	143.1 (144)	r
Yellow	$\sqrt{80}$	¥	$\sqrt{68}$	77.5 (72)	102.5 (108)	R
Cyan	$\sqrt{89}$, ¥	$\sqrt{68}$	44.0 (45)	136.0 (135)	ℓ
Blue	$\sqrt{89}$, ¥	$\sqrt{80}$	84.6 (90)	95.4 (90)	Q
Magenta	$\sqrt{68}$	=	$\sqrt{68}$	61.9 (60)	118.1 (120)	T
Green	$\sqrt{89}$	=	$\sqrt{89}$	26.0	154.0	с

is crosswise supported by the previously mentioned analogy of an (in)commensurate modulation, since the interstitial points to the pseudo-decagonal clusters are subject to the largest distortions in Hammersley tilings of larger size (*cf.* Fig. 2*b*).

Both the Penrose and the Ammann–Beenker rhombs intermingle in a peculiar way, such that it becomes difficult to classify the tiling either as more of the Penrose or more of the Ammann–Beenker type, although the presence of pseudodecagonal clusters favours the connection to the Penrose tiling.

Of course there are obvious differences. A single decagon is constituted of three thin and four thick rhomb(oid)s in addition to one equilateral triangular rhomb in the centre. In particular, two of the thin rhombs account only for one half of the area of the decagon. The combination of tiles thus does not obey the matching rules applicable in the case of a quasiperiodic Penrose tiling. Of the eight valid vertex configurations to be found within a Penrose tiling, only one, the Kconfiguration, describes a quadruple of tiles around a single vertex. The K-configuration is given by the arrangement of one thin (r) together with three thick (R) rhombs, according to the scheme $r_0 R_a^3$, where the subscripts a and o denote the orientation of the tiles towards the common vertex referring to either an acute or obtuse angle, respectively. The combination $r_a R_o^3$ would also be possible regarding the sum of its angles, $1 \times 36^{\circ} + 3 \times 108^{\circ} = 360^{\circ}$, but is not a valid vertex configuration in a Penrose tiling. The decagon of the Hammersley tiling of Fig. 2(a) encloses four vertices, of which two each form a vertex configuration of $r_0 R_a^3$ and $r_a R_o^3$, respectively. A second difference is due to the fact that within a rhombic Penrose tiling two types of decagons may occur, whereas the tiling under consideration in this work is constituted of a single type of decagonal cluster. Yet another difference concerns the frequency of occurrence of thin and thick tiles. Here, the ratio of the number of thick and thin Penrose-like rhomb(oid)s is given by $R/r = 1 < \varphi$, whereas in a quasiperiodic Penrose pattern this would be $R/r = \varphi$, where $\varphi = (1 + \sqrt{5})/2$ is the golden mean. A similar relation holds for the quasiperiodic Ammann-Beenker tiling for which the ratio of squares to lozenges is given as $Q/\ell = 1/\sqrt{2}$ in contrast to a larger ratio $Q/\ell = 2$ in our case.

In any way the relation is closer to the binary Penrose tiling, or *binary tiling* for short, characterized by matching rules that differ from those of the original Penrose tiling (Lançon & Billard, 1988, 1993). Similar Penrose-like tilings were studied by Widom *et al.* (1987). In a more general way the Hammersley tiling strongly resembles a combination of binary (canonical) substitution tilings of Penrose-type and fivefold symmetry (Godrèche & Lançon, 1992) and of Ammann-Beenker type and eightfold symmetry (Harriss & Lamb, 2004; *cf.* in particular tiling M_3), respectively. Eventually, a further generalization to binary tilings with *n*-fold symmetry, where *n* is odd, is possible (Lançon & Billard, 1993).

The overall arrangement of decagonal clusters evokes the impression of a perfect square arrangement, and the pattern may certainly count as a near-miss in this respect too, but its construction from a permutation impedes this type of regularity (a wedge indicates the associated horizontal offset in Fig. 3). However, the patch is symmetric with respect to its diagonals, thereby enforcing all parallelograms situated along these diagonals to be rhombs rather than rhomboids. By analogy the pattern shown in Fig. 3, and certainly any of its enlarged successors, could be conceived as a finitely extended periodic approximant to a quasiperiodic structure. Furthermore, a bit-truncated version of the tiling indeed exhibits the more regular square arrangement of decagonal clusters (*cf.* Fig. 4 in Appendix A2).

3. Quasirandomness at large

Aside from the characterization of quasirandom number sequences, two-dimensional point sets and planar tilings presented in §2, we draw the reader's attention to the intimately related concept of *discrepancy measures* and consider its possible implications for crystallography.

3.1. Quantifying order in space

In order to quantitatively assess the uniformity of a given point distribution, the concept of *discrepancy* as an appropriate measure is introduced. The (local) discrepancy

$$D(P,S) = \left| \frac{N_S}{N} - \frac{V_S}{V} \right| \tag{1}$$

of a given set *P* of *N* points is the measure specifying the difference between the actual number density of a subset of points, N_S/N , within a chosen subinterval *S*, with respect to the expected one, according to the relative spatial extension (area, volume *etc.* depending on spatial dimension) of the subinterval, V_S/V , and perfect uniformity provided. *V* is commonly represented by the *s*-dimensional unit interval $[0, 1)^s$, thus V = 1. A global discrepancy may then be defined to be the least upper bound, *i.e.* the supremum, with respect to all possible discrete subintervals and their local discrepancies under consideration.

Some alternative discrepancy measures exist, such as the socalled star discrepancy D^* , which may have certain advantages in practical applications, in particular regarding the ease of their calculation. Matoušek (2010) gives a comprehensive survey of these and other issues of geometric discrepancy theory as do Kuipers & Niederreiter (1974; ch. 2) in the context of uniform distribution of sequences; see also the concise discussion by Entacher (2005).

In general, discrepancy theory is concerned with the deviations from uniformity in a way similar to Ramsey theory which studies the conditions under which some predefined order must appear. Discrepancy measures complement other established methods of detecting and quantitatively judging order, such as nearest-neighbour statistics, radial distribution functions or Fourier methods (Illian *et al.*, 2008).

Irrespective of the chosen definition, any discrepancy measure is the lower the more uniform a distribution of points becomes (and zero in the limit of perfect uniformity). Notably, a point lattice, intuitively appearing most uniform, does not necessarily possess a low discrepancy, owing to its arrangement of points being too regular. A number of methods exist for the construction of low-discrepancy sequences, among them, in particular, the one devised by van der Corput (*cf.* Appendix A1).

3.2. Quasirandomness in nature?

Crystal structures, commonly conceived to be prime examples for the ordered arrangements of atoms, governed by symmetry, may rather be representatives of quasi- and pseudorandom principles.

The occurrence of approximate decagonal clusters extending into an almost periodic variant of the Penrose tiling made of thin and thick rhombs in a rather simple numbertheoretic context of quasirandom low-discrepancy sequences may be the result of a mere coincidence. However, although the generation of the quasirandom van der Corput set cannot be in any way compared with the formation of an actual quasicrystal, the underlying mechanisms may be nevertheless quite similar. Most often quasicrystals are rapidly solidified from their comparatively disordered molten state, in which the atoms may be distributed in a rather random fashion. At the same time crystallization occurs under severe physical constraints, e.g. regarding the mutual self-avoidance of atoms due to classical and quantum mechanical repulsion principles, which set a lower bound for interatomic distances and thereby impose a kind of uniform distribution similar to a quasirandom point set. The entropy-driven randomness in the aggregation of atoms is accompanied by enthalpy-driven correlations in their spatial arrangement, which in turn lends the whole process a qualitative quasirandom characterization. These arguments should be even more persuasive considering the case of amorphous solids, where the arrangement of atoms is regarded as 'random'; certainly not truly random as in the sense of Fig. 1(d), but rather quasirandom instead.

Thus, one may think about *quasirandomness* as a governing factor in the formation of crystal structures, actual realisations of abstract, uniformly discrete and relatively dense point sets, similar in scope but distinct in its consequences to symmetry as a fundamental ordering principle.

In particular, it seems possible that the *discrepancy* of a crystal structure (*cf.* $\S3.1$), *i.e.* the discrepancy of the three-

dimensional point set derived from its atomic coordinates, is not only a measure of the uniformity in the distribution of atoms but itself a variable under optimization during the formation of the crystal structure. It may also be that the tendency of atoms to arrange themselves into structures of highest symmetry is challenged by a tendency to be distributed in a most uniform way, *i.e.* to minimize geometric discrepancy or, in other terms, to maximize self-avoidance, instead. Crystallization then may be conceived as a complex optimization process regarding (de)localized chemical bonding and mutual (discrepancy-driven) self-avoidance, much as in Schopenhauer's parable *Die Stachelschweine* (The Porcupines), loosely restrained by more global factors like packing density and symmetry.

An example of this may be given by the structure of β -manganese (Lidin & Fredrickson, 2012), which proves to be a near-miss to a three-dimensional metrically cubic permutation structure, however of low triclinic symmetry (Hornfeck, 2012). Owing to slight distortions from the ideal permutation structure, with the more uniform spatial distribution of atoms, the actual crystal structure attains cubic symmetry instead.

Eventually this may also give an explanation for why quasicrystalline states of matter could be in some cases superior to crystalline ones and why quasicrystalline states often exhibit a residual disorder regarding a potential state of higher symmetry. Arguments based on discrepancy theory should facilitate an analysis in quantitative terms.

In any way, the rather subtle semantic difference between pseudo- and quasirandomness may challenge our fundamental understanding of matter and emphasize the continuous interplay between all kinds of possible states of short- and long-range order.

3.3. Combinatorial crystallography

The aforementioned observations call for some qualitative assessments, focusing on some ideas of how to take a distinct look at crystal structures.

We would like to emphasize the arithmetical properties of crystal structures, i.e. their quantitative details beyond the mere listing of atomic coordinates or the calculation of interatomic distances and angles and in contrast to the qualitative classification regarding symmetry. We deem it useful to think of crystal structures in a context of what one may call crystallographic information theory, i.e. as spatial codes, as actual expressions of chemical interactions, representing a case in which the messenger is identical to the message [a similar, even more comprehensive, perspective is developed in a recent essay of Cartwright & Mackay (2012)]. One may even think about the algorithmic generation of crystal structures similar to the Lindenmayer systems of theoretical biology (Lindenmayer, 1968). In the latter case the alphabet would consist of a set of elements, including vacancies as null-elements, a set of substitution rules and an initial string from which the iterative generation of the structure proceeds, quite similar to concepts used in the description of homologous series of crystal structures or quasicrystals.

In this context we would like to draw attention to the arithmetic approaches of Nardone (2006) and Indelicato (2013, and references therein) in their studies of *multilattices* (where a multilattice is defined as a finite union of translates of a given simple lattice). Their approach extends the method on which the classification of simple lattices into Bravais types is based in order to facilitate a finer classification scheme for crystal structures than the classical one using space-group types alone.

Arithmetical methods based on the generation of random numbers seem to be another approach by which structures, *i.e.* point patterns, can be encoded in a most general way. Symmetry *emerges* quite naturally from the cycle structure of the associated permutations, for the pseudorandom numbers generated by the multiplicative congruential method introduced by Lehmer (1949) as well as for the quasirandom ones generated by the bit-reversal method of van der Corput (1935).

4. Conclusion

Quasirandom point sets constructed by the bit-reversal method of van der Corput exhibit certain features of quasicrystals, notably pseudo-decagonal short-range order. Their associated planar tilings, termed Hammersley tilings H_n , are almost completely constituted of near-miss rhombic tiles closely approximating their ideal counterparts from the rhombic tilings of Penrose and Ammann-Beenker. Furthermore, these tilings resemble the semi-regular ones composed of squares and rhombs of Baake & Grimm (2012), termed planar σ -phases and showing pseudo-symmetrical intensity distributions in their diffraction patterns with apparent symmetries characteristic of axial quasicrystals. For each of the considered quasirandom tilings there is a corresponding description by means of bit-reversal permutations, in many kinds similar to permutations arising from multiplicative congruential generators, which are known to generate pseudorandom numbers on their part.

In conjunction with the discussion of quasirandom sequences, point sets and tilings, the concept of discrepancy was evoked, discussing its ability to quantify distinct states of order in between the limiting states of purely crystalline and purely amorphous condensed matter. For instance, it seems tempting to explore discrepancy measures applied to crystal and quasicrystal structures, but in particular to the structures of disordered solids including what is known about structures representing the amorphous state.

Random point sets have been studied from the point of view of their diffraction (Baake & Kösters, 2011; Baake & Grimm, 2012), but it seems equally fruitful to analyse their real-space properties. In either view a definition of long-range order is still missing, regardless of its definition in real or dual space. From the analysis presented in this note one may conclude that the study of pseudo- and quasirandom point sets and their related permutations may add something to our understanding of the arrangement of atoms in condensed matter. To us it appears that a vast amount of knowledge already collected in the fields of discrete mathematics and computer science merely awaits its translation into and application within a crystallographic context.

APPENDIX A

van der Corput sequences and Hammersley sets

A1. van der Corput's construction

A sequence of given length M is constructed from the set of M successive non-negative integers $N_i \in \{0, 1, \dots, M-1\} = \mathbb{Z}/M\mathbb{Z}$ by first expanding them into their respective binary representation

$$N_i = \sum_{n=0}^{m-1} b_n(N_i) 2^n = b_{m-1} \dots b_1 b_0$$
(2)

followed by a bitwise reversal

$$N_i^* = \sum_{n=0}^{m-1} \frac{b_n(N_i)}{2^{n+1}} = 0.b_0 b_1 \dots b_{m-1}$$
(3)

and their final back-conversion into the corresponding decimal numbers, thereby yielding the base 2 van der Corput sequence.

A two-dimensional point set is then constructed by a *pairwise combination* of the initial and final numbers of the bitreversal process, $p_i = (N_i/M, N_i^*)$, the former ones taken as fractions of the common denominator M,

$$N_{i}/M \to \frac{b_{m-1} \, 2^{m-1} + \ldots + b_{0} \, 2^{0}}{2^{m}} = \frac{b_{m-1}}{2^{1}} + \ldots + \frac{b_{0}}{2^{m}},$$
$$N_{i}^{*} \to \frac{b_{0}}{2^{1}} + \ldots + \frac{b_{m-1}}{2^{m}}.$$
(4)

Here, the bit-reversal process becomes obvious.

A decagonal pattern first appears for a point set of size $M = 2^4 = 16$ for which the coordinate values are listed in Table 2, illustrating the bit-reversal process in detail.

The method of van der Corput is naturally generalized to distinct bases other than 2 as well as to arbitrary dimension s > 1 (*Halton sets*). In addition, the distribution properties of the sequence may be altered by a bit-truncation process accompanying the bit reversal (*Hammersley sets*). The latter process is interesting, since it introduces a kind of finite periodicity ultimately leading to a lattice-ordered state of square symmetry.

A2. Bit-truncated Hammersley sets

The pattern of 64 points depicted in Fig. 1(c) is an example of a so-called (t, m, s)-net, in particular representing the (0, 6, 2) case, where 2^m defines the size M of the set (*i.e.* the number of points; here, m = 6, M = 64), s denotes the dimension (here, two-dimensional) and the integer t would define an additional intermediate truncation of t bits, taking place after the bit-reversal process but prior to the plotting of coordinate values. The (t, m, s)-nets of (0, m, 2)-type are known as Hammersley sets, thus generalizing the principle underneath the van der Corput sequence to higher dimensions. The prime example of a low-discrepancy sequence, the

Table 2

Point coordinates $p_i = (N_i/M, N_i^*)$ of the van der Corput set of size M = 16 and their generation via bit reversal.

The subscripts 2 and 10 denote the base and $(N_i^*)_{10}$ is given in multiples of M^{-1} in order to emphasize the permutation character of the process [*cf.* equation (5)].

$(N_i)_{10}$	$(N_i)_2$	$(N_{i}^{*})_{2}$	$(N_{i}^{*})_{10}$	$(N_i)_{10}$	$(N_i)_2$	$(N_{i}^{*})_{2}$	$(N_{i}^{*})_{10}$
0	0	0.0	0	8	1000	0.0001	1
1	1	0.1	8	9	1001	0.1001	9
2	10	0.01	4	10	1010	0.0101	5
3	11	0.11	12	11	1011	0.1101	13
4	100	0.001	2	12	1100	0.0011	3
5	101	0.101	10	13	1101	0.1011	11
6	110	0.011	6	14	1110	0.0111	7
7	111	0.111	14	15	1111	0.1111	15

van der Corput sequence, may then be conceived as a special case, namely as the (t, s)-sequence with t = 0 and s = 1. In some sense the Hammersley point set thus transcends the one-dimensional van der Corput sequence into a two-dimensional pattern.

Whereas the (0, 6, 2)-net of Fig. 1(c) or Fig. 4(a) is only a *near miss* to a pattern with a square arrangement of clusters, the bit-truncated (1, 6, 2)-net exhibits perfect repetition, as depicted in Fig. 4(b). Bit truncation thus destroys the permutation structure but at the same time enforces a (finite) periodicity compared with the original set. Higher degrees of bit truncation with t > 1 yield subdivisions of increasingly smaller repetition period, with a minimum at t = m/2; henceforth square lattice patterns prevail, thereby illustrating the information loss accompanying each bit-truncation step. Bit truncation to the highest degree, *i.e.* t = m, is identical to the projection of the whole set of numbers to a single point at (0, 0).

The effect of the bit-truncation process may be exemplified by comparison of the circled regions in the upper row of Fig. 4 which contain a single point each. Whereas for the t = 0pattern both points are located at distinct positions with respect to a common horizontal axis, their position coincides for the t = 1 pattern. If one superimposes a grid, the grid lines will be separated with a spatial resolution of $1/2^{m-t}$ and each grid line will contain 2^t points. For the pattern with t = 0 the spacing of the grid is minimal, here 1/64, thus maximizing its resolution, but each grid line contains only one point.

Thus, for low values of t, bit truncation 'acts' as a small distortion yielding finite analogues to an infinitely extended pattern with translation periodicity, whereas high values of t eventually enforce the appearance of a primitive square lattice. Thereby, the bit-truncation process creates a quasi-continuous transformation between distinctly ordered states, with the most uniform (t = 0) and the most regular (t = m) distribution of points at its limits.

A detailed exposition of the aforementioned is given by Entacher (2005), illustrating the use of the *Mathematica* software in the construction and statistical analysis of quasi- and pseudorandom point sets (Hammersley and Halton point sets as well as ones generated by the action of linear and non-linear congruential generators, the former including MCGs).



Figure 4

Hammersley point sets (t, 6, 2) for distinct bit truncations t with dashed lines denoting unit-cell boundaries. For t = 1 one has four copies of the Hammersley point set (0, 4, 2), *i.e.* colloquially $(1, 6, 2) = 4 \times (0, 4, 2)$. Note the onset of a series of square lattices, beginning from t = 3(generally t = m/2) onwards. At t = 6 everything collapses into a single point at (0, 0), corresponding to a full truncation of all bits, thereby completing the circle of decreasing and increasing number of unit cells, 1, 4, 16, 32, 16, 4, 1, within the fixed frame defined by the t = 0configuration.

In addition, (t, s)-sequences and (t, m, s)-nets are treated in various contexts in the monographs of Niederreiter (1992) and Hellekalek & Larcher (1998).

A3. Bit-reversal permutations

The coordinates listed in Table 2 illustrate the bit-reversal process yielding a base-two van der Corput sequence of size $M = 2^4 = 16$. Here, the emphasis is laid on the associated bit-reversal permutations whose existence arises from the fact that the size of the sequence is given by an (even) power of two. This guarantees that the set of M integers $N_i^*M \in \mathbb{Z}/M\mathbb{Z}$ is in a one-to-one correspondence with the set $N_i \in \mathbb{Z}/M\mathbb{Z}$, defining the permutation (in cycle notation)

 $(0)(1 \ 8)(2 \ 4)(3 \ 12)(5 \ 10)(6)(7 \ 14)(9)(11 \ 13)(15),$ (5)

solely consisting of cycles of length $\ell = 1$ (fixpoints) and $\ell = 2$ (transpositions). The same argument holds for all point sets of size $M = 2^{2n+2}$ corresponding to a Hammersley tiling H_n . Accordingly, the set for $M = 2^6 = 64$ may be represented as

(0)(1 32)(2 16)(3 48)(4 8)(5 40)(6 24)(7 56)(9 36) (10 20)(11 52)(12)(13 44)(14 28)(15 60)(17 34)(18) (19 50)(21 42)(22 26)(23 58)(25 38)(27 54)(29 46) (30)(31 62)(33)(35 49)(37 41)(39 57)(43 53)(45) (47 61)(51)(55 59)(63).

Owing to a spatial repetition this may be noted much shorter,

 $(0)(2\ 16)(4\ 8)(6\ 24)(10\ 20)(12)(14\ 28)(18)(22\ 26)(30)$

+ (1 32); +(32 1); +(33).

Dividing all entries by two (disregarding repetition) yields the permutation according to the van der Corput sequence of length M = 16 [*cf.* equation (5)]. In a similar way the van der Corput sequence of length M = 256 is given by the permutation described in equation (5) *via* a multiplication by four:

(0)(4 32)(8 16)(12 48)(20 40)(24)(28 56)(36)(44 52)(60)

in addition to a set of 15 translations, *e.g.* along $+(1\ 128)$, emphasizing the hierarchical construction principle of these permutations as well as their associated point sets and tilings.

In general, any permutation corresponding to a Hammersley tiling H_n of size M in base two consists solely of m_1 fixpoints and m_2 transpositions, which are trivially related according to $M = 2m_2 + m_1 = 2^{2n+2}$. The number of fixpoints, including the origin (0), is given as $m_1 = \sqrt{M}$, and consequently $m_2 = (m_1^2 - m_1)/2$.

A4. Generalized Hammersley tilings

At the lowest hierarchical level the Hammersley tiling H_1 comprises a single pseudo-decagonal building block containing nine tiles of three distinct types [eight tiles surrounding a central one, according to a one-line formula $T(Rr)_4$, where each tile is represented by its one-letter symbol] encompassing 16 vertices,

				r	R	r	Q	r	R	r
				R	Т	R	ℓ	R	Т	R
r	R	r		r	R	r	Q	r	R	r
R	Т	R	\rightarrow	Q	ℓ	Q	С	Q	ℓ	Q
r	R	r		r	R	r	Q	r	R	r
				R	Т	R	ℓ	R	Т	R
				r	R	r	Q	r	R	r
	H_1						H_2			

The next level of spatial extension, represented by the tiling H_2 , repeats the structural motif of the former one, however, by means of adding new distinct types of tiles at the barycentre c and in between the pseudo-decagonal clusters (Q, ℓ) . Thus,

Table 3

First representatives and general formulas of the Hammersley tilings H_n counting the number of vertices, tiles, pseudo-decagonal clusters and intercluster barycentres, as well as cluster packing fraction ρ .

The value of 9/16 for the cluster packing fraction describes the limit for an infinitely extended tiling H_{∞} .

Tiling	Filing Vertices Tiles		Cluster	Barycentres	Q
H_1	16	9	1	0	1
H_2	64	49	4	1	0.7347
H_{3}	256	225	16	9	0.6400
H_{4}	1024	961	64	49	0.5994
H_{5}	4096	3969	256	225	0.5805
H_6	16384	16129	1024	961	0.5714
H_7	65536	65025	4096	3969	0.5669
:			:		÷
H_n	$(2^{n+1})^2$	$(2^{n+1}-1)^2$	$(2^{n-1})^2$	$(2^{n-1}-1)^2$	0.5625
'n	$= 2^{2n+2}$		$= 2^{2n-2}$		= 9/16

the one-line notation for H_2 may be given as $c(Q_2\ell)_4\{T(Rr)_4\}_4$. A comparison with the H_1 tiling shows the *iterative* and *creative* nature while proceeding from one level of hierarchy to the next. For every step $H_n \mapsto H_{n+1}$ the finite patch H_n is repeated in a fourfold arrangement surrounding a newly created barycentre separated by four newly created single bands of tiles. This creation of new kinds of tiles renders a general formula including the frequency of occurrence of special rhomb(oid)s less useful than simply counting the number of vertices, tiles, clusters and barycentres as listed in Table 3.

A5. Iterated maps and the van der Corput sequence

Multiplicative congruential generators (MCGs) were shown to be connected to iterated maps [see Hornfeck (2012) for details and references]. The same is true for the bit-reversal process related to the van der Corput sequence, which is generated as the orbit of the origin under the dyadic von Neumann–Kakutani transformation (Grabner *et al.*, 2012). Twelve initial steps of the orbit are depicted by arrows in Fig. 5.



Figure 5

Iterated map based on the dyadic von Neumann-Kakutani transformation.

Both the start and the end point of the orbit, the latter located at the diagonal y = x, are marked by open circles. As is the case for MCGs, the iterated map based on the von Neumann–Kakutani transformation is defined by a piecewise linear function constructed by a splitting and stacking procedure.

APPENDIX **B**

Phase-space sampling by pseudo- and quasirandom numbers in structure solution

A great part of the successful application of *direct methods* for the solution of crystal structures may be traced back to the efficient generation of a finite starting set of random phases. The practical task consists of finding an efficient sampling of the highly multi-dimensional phase space, in particular if one is considering the general non-centrosymmetric case.

For this purpose different strategies have been devised, among them the so-called *magic integer method*, introduced by White & Woolfson (1975). The main idea of their approach can be traced back to ideas first discussed by Woolfson (1954) and later justified by Good (1954, 1959) regarding the determination of centrosymmetric structures by means of sign permutations. The method was furthermore extended and given a theoretical foundation in a series of articles regarding *The Application of Phase Relationships to Complex Structures*, especially by Declercq *et al.* (1975) and Main (1977, 1978).

The method is based on the approximative representation of a given set of phases $0 \le \phi_i < 1$ by a set of *n* equations,

$$\phi_i \equiv m_i x \pmod{1}, \qquad i = 1, \dots, n, \tag{6}$$

where $0 \le x < 1$ is a fixed parameter to be determined and m_i constitute the set of *n* magic integers. By this means it is possible to represent a large number of phases in terms of a comparatively small number of parameters.

Furthermore, the representation of single phase sets by means of magic integers can be extended to encompass relationships between phases. Given, for instance, a triplet relationship

$$\phi_i + \phi_i + \phi_k + \pi \equiv 0 \pmod{1} \tag{7}$$

with phases ϕ_i represented by the same basic set of magic integers m_i in several variables x_i . For instance, in the case of three variables x, y, z replacing the phases by their magic integer representation and *combining like terms*, *i.e.* making a summation over distinct magic integer coefficients for the same variable, leads to a relationship

$$Hx + Ky + Lz + b \equiv 0 \pmod{1},\tag{8}$$

which may be used, in the course of a Fourier synthesis, to determine a refined phase set which after additional steps eventually leads to the solution of the structure. The main advantage lies in the efficient and simultaneous representation of a primary set of single phases and phase relationships, which are then used to obtain a secondary set of phases and thereby extending them until a solution is found. Although the approach is of an approximative nature, the large number of single phases and phase relationships guarantee the robustness of the algorithm with respect to inevitable errors in the accuracy of determining the correct phases.

Magic integers thus may be used both (and in fact simultaneously) for phase-space exploration and phase extension given a set of already known phases as well as for the generation of a starting set of phases and their subsequent permutation given a set of yet unknown phases (Declercq *et al.*, 1975).

The magic integer method allows for a much more efficient phase-space sampling than the earlier concept of quadrant permutation. An explanation for this is given by analogy: the magic integer method represents the case of a point set generated by the multiplicative congruential method, the generated numbers thus are pseudorandom and therefore superior to the case of quadrant permutation, which represents the lattice-ordered case [*cf.* Fig. 6 of Main (1978)]. To the best of our knowledge, no one has yet explored possible applications of quasirandom numbers in the application of direct methods, although they should, in principle, allow for an even more uniform sampling of phase space than is the case for pseudorandom numbers.

The set of congruence relations defined by equation (6) describes the parametric equation of a straight line (subject to periodic boundary conditions) within an *n*-dimensional phase space in a similar way as was independently described by Hornfeck (2012) for the real-space description of crystal structures by means of the multiplicative congruential method. This connection between real- and dual-space concepts may seem surprising; however, the underlying sampling tasks are of an analogous nature. As yet it seems unclear whether both approaches can be combined in such a way that would improve crystal structure solution by means of direct methods.

Many thanks to Philipp Kuhn for comments on the manuscript, to Sven Binder for his subversive mindset, and to Markus Blümner for inspiration.

References

- Baake, M. & Grimm, U. (2012). Chem. Soc. Rev. 41, 6821-6843.
- Baake, M. & Kösters, H. (2011). Philos. Mag. 91, 2671-2679.
- Cartwright, J. H. E. & Mackay, A. L. (2012). *Philos. Trans. R. Soc. A*, **370**, 2807–2822.
- Corput, J. G. van der (1935). Proc. Ned. Akad. Wet. 38, 813-821.
- Declercq, J. P., Germain, G. & Woolfson, M. M. (1975). Acta Cryst. A**31**, 367–372.

- Entacher, K. (2005). *Math. J.* **9**(3) (http://www.mathematica-journal. com/issue/v9i3/UniformDistribution.html).
- Godrèche, C. & Lançon, F. (1992). J. Phys. I Fr. 2, 207-220.
- Good, I. J. (1954). Acta Cryst. 7, 603-604.
- Good, I. J. (1959). Acta Cryst. 12, 824-825.
- Grabner, P. J., Hellekalek, P. & Liardet, P. (2012). Unif. Distrib. Theory, 7, 11-70.
- Grünbaum, B. & Shephard, G. C. (1987). *Tilings and Patterns*. New York: W. H. Freeman.
- Hammersley, J. M. & Handscomb, D. C. (1975). Monte Carlo Methods. Methuen's Monographs on Applied Propability and Statistics. London: Methuen.
- Harriss, E. O. & Lamb, J. S. W. (2004). *Theor. Comput. Sci.* **319**, 241–279.
- Hayes, B. (2011). Am. Sci. 99, 282–287 (see also http://bit-player.org/ 2011/a-slight-discrepancy).
- Hellekalek, P. & Larcher, G. (1998). Editors. Random and Quasi-Random Point Sets. Lecture Notes in Statistics 138. New York/ Berlin/Heidelberg: Springer.
- Hornfeck, W. (2012). Acta Cryst. A68, 167-180.
- Hornfeck, W. & Harbrecht, B. (2009). Acta Cryst. A65, 532-542.
- Illian, J., Penttinen, A., Stoyan, H. & Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns. Statistics in Practice.* Chichester: John Wiley and Sons.
- Indelicato, G. (2013). Acta Cryst. A69, 63-74.
- Kuipers, L. & Niederreiter, H. (1974). Uniform Distribution of Sequences. New York: John Wiley and Sons.
- Lançon, F. & Billard, L. (1988). J. Phys. Fr. 49, 249-256.
- Lançon, F. & Billard, L. (1993). Phase Transit. 44, 37-46.
- Lehmer, D. H. (1949). Ann. Comput. Lab. Harvard Univ. 26, 1951.
- Lidin, S. & Fredrickson, D. (2012). Symmetry, 4, 537-544.
- Lifshitz, R. (2002). J. Alloys Compd. 342, 186-190.
- Lindenmayer, A. (1968). J. Theor. Biol. 18, 280-299.
- Main, P. (1977). Acta Cryst. A33, 750-757.
- Main, P. (1978). Acta Cryst. A34, 31-38.
- Marsaglia, G. (1968). Proc. Natl Acad. Sci. 61, 25-28.
- Matoušek, J. (2010). Geometric Discrepancy An Illustrated Guide. Algorithms and Combinatorics 18. New York/Berlin/Heidelberg: Springer.
- Nardone, M. (2006). J. Math. Phys. 47, 033509.
- Neumann, J. von (1951). In Monte Carlo Methods, edited by A. S. Householder, G. E. Forsythe & H. H. Germond, National Bureau of Standards Applied Mathematics Series 12, pp. 36–38. Washington: US Government Printing Office.
- Niederreiter, H. (1992). Random Number Generation and Quasi-Monte Carlo Methods. Philadelphia: Society for Industrial and Applied Mathematics.
- Sire, C., Mosseri, R. & Sadoc, J.-F. (1989). J. Phys. Fr. 50, 3463-3476.
- White, P. S. & Woolfson, M. M. (1975). Acta Cryst. A31, 53-56.
- Widom, M., Strandburg, K. J. & Swendsen, R. H. (1987). Phys. Rev. Lett. 58, 706–709.
- Woolfson, M. M. (1954). Acta Cryst. 7, 65-67.