

Algebraic Description of Coordination Sequences and Exact Topological Densities for Zeolites

R. W. GROSSE-KUNSTLEVE,^{a*†} G. O. BRUNNER^a AND N. J. A. SLOANE^b

^aLaboratory of Crystallography, ETH Zentrum, CH-8092 Zurich, Switzerland, and ^bMathematical Sciences Research Center, AT&T Research, Murray Hill, New Jersey 07974, USA. E-mail: rwgk@laplace.csb.yale.edu

(Received 27 February 1996; accepted 6 June 1996)

Abstract

Coordination sequences (CS) have been calculated for all approved zeolite topologies, all dense SiO₂ polymorphs and 16 selected non-tetrahedral structures and the algebraic structure of these CS's has been analyzed. Two algebraic descriptions of coordination sequences are presented. One description uses periodic sets of quadratic equations and is already established in the literature. The second description employs generating functions, which are well known in combinatorics but are used here for the first time in connection with coordination sequences. The algebraic analysis based on generating functions turns out to be more powerful than the other approach. Based on the algebraic analyses, exact topological densities are derived and tabulated for all the structures investigated. In addition, 'n-dimensional sodalite' is observed to have an especially simple n-dimensional graph.

1. Introduction

The notion of coordination sequence (CS) was formally introduced by Brunner & Laves (1971) in order to investigate the topological identity of frameworks and of atomic positions within a framework. The CS is a number sequence in which the *k*th term is the number of atoms in 'shell' *k* that are bonded to atoms in 'shell' *k* – 1. Shell 0 consists of a single atom and the number of atoms in the first shell is the conventional coordination number.

The CS is now routinely used to characterize crystallographic structures (Meier & Möck, 1979; *Atlas of Zeolite Structure Types*, 1992, 1996; Fischer, 1973, 1974) and even higher-dimensional sphere packings (Conway & Sloane, 1995, 1996). Other applications are to the determination of topological density, which can be obtained from the partial sums of terms in the CS. This density is correlated with other parameters such as lattice energy (Akpoyiaye & Price, 1989), the distribution of particular elements (Herrero, 1993) and catalytic activity (Barthomeuf, 1993), and is useful for predicting properties of synthetic zeolites (Brunner, 1979).

In an abstract sense, the CS describes the growth of a crystal and was therefore initially called the 'growth series' (Brunner & Laves, 1971). Previous investigations (Brunner, 1979; Herrero, 1994; Schumacher, 1994; O'Keefe, 1991a) have shown that in some cases the terms in the CS increase quadratically with *k* but up to now investigations have been restricted to specific examples. In view of the increasing applications of the CS, up to 2000 terms have now been calculated for all approved zeolite topologies as well as for all dense SiO₂ polymorphs and 16 selected non-tetrahedral structures, and the algebraic structure of these CS's has been analyzed.

2. The crystal structures investigated

The majority of the zeolite structures investigated in this work are listed in the *Atlas of Zeolite Structure Types* (1992, 1996).‡ All the zeolite topologies will be referred to by their three-letter codes. For example, melanophlogite has code MEP. Coordination sequences for structures with more than one crystallographic site are referred to by the three-letter code followed by the site label of the *Atlas*. For example, the zeolite mazzite has two distinct sites (both tetrahedrally coordinated atoms), which are denoted by MAZ T1 and MAZ T2.

The data for the 12 SiO₂ polymorphs were taken from the Inorganic Crystal Structure Database (1986–1995). Table 1 lists the mineral names together with the ICSD collection codes.

For a comparison, 16 non-tetrahedral structure types having many chemical representatives have also been investigated. Table 2 gives the search codes for the corresponding entries in the *Gmelin Handbook of Inorganic and Organometallic Chemistry* (TYPIX, 1994).

The coordination numbers of the zeolites and the dense SiO₂ polymorphs are always 4, with the exception of the six interrupted zeolite frameworks (indicated by a dash preceding the three-letter code), which have a three- or twofold coordination for one of the sites. For the non-tetrahedral structures, Table 2 also gives the bond-length

† Present address: Yale University, New Haven, CT 06520-8114, USA.

‡ The *Atlas of Zeolite Structure Types* is available on the World Wide Web at <http://www.iza-sc.ethz.ch/IZA-SCI/>.

Table 1. *ICSD collection codes*

Mineral	ICSD collection code
Banalsite	4447
Coesite	18112
Cordierite	30947
Cristobalite	9327
Feldspar	100182
Keatite	34889
Milarite	71046
Moganite	67669
Paracelsian	24690
Quartz	31048
Scapolite	9502
Tridymite	29343

Table 2. *TYPIX search codes*

Formula	TYPIX search code	Bond length limit (Å)	Coordination numbers
CaF ₂ (2)	(225) cF12	2.5	4, 8
CaF ₂ (1)	(225) cF12	3.5	8, 10
NaCl	(225) cF8	3.0	6
FeS ₂ -marcasite	(58) oP6	2.7	4, 6
FeS ₂ -pyrite	(205) cP12	2.7	4, 6
NiAs (2)	(194) hP4	2.5	6, 6
NiAs (1)	(194) hP4	2.7	6, 8
Cu	(225) cF4	2.7	12
Mg	(194) hP2	3.5	12
W (2)	(229) cI2	3.3	14
W (1)	(229) cI2	2.8	8
α-Nd	(194) hP4	4.0	12
Ni ₂ In	(194) hP6	3.3	11, 14
α-Mn	(217) cI58	3.5	12, 13, 16
Cr ₃ Si	(223) cP8	3.5	12, 14
α-CrFe	(136) tP30	3.5	12, 14, 15
MgZn ₂	(194) hP12	3.5	12, 16
MgCu ₂	(227) cF24	3.5	12, 16
MgNi ₂	(194) hP24	3.5	12, 16

limits and the resulting coordination numbers that were used in the calculations. For each of CaF₂, NiAs and W, two different parameters were considered, leading to two different coordination numbers.

3. Primary determination of coordination sequence: a highly optimized node-counting algorithm

In order to calculate the coordination sequence, the crystal structure, *i.e.* assembly of atoms, has to be abstracted to a mathematical topology (or graph) with nodes and certain bonds between nodes. In the case of zeolites and dense SiO₂ phases, the tetrahedral positions are taken to be the nodes, and the bridging framework of O atoms is replaced by bonds. For other classes of materials, *e.g.* metals or intermetallic phases, all atoms represent nodes, and bonds are created in an appropriate neighborhood of each atom (see, for example, Brunner & Laves, 1971).

The CS determination algorithm used here can be described as a node-counting or *coordination-shell algorithm*. The algorithm is started (with $k = 0$) by selecting an initial node. At the next step ($k = 1$), all nodes

bonded to the initial node are determined. For $k \geq 2$, all characteristics of the algorithm become evident: those nodes that are bonded to the 'new nodes of the previous step ($k - 1$)' but have not been counted before are counted. This means that three sets of nodes for three topological distances (*i.e.* three coordination shells) have to be maintained: the middle ($k - 1$) nodes, whose bonds are followed to determine the next (k) nodes, and the previous ($k - 2$) nodes, to know which of the nodes bonded to the middle nodes have already been counted. The innermost shells with $k < k_{\text{next}} - 2$ are not needed and can be deleted since the nodes counted before are fully surrounded by shell $k - 2$. In this way, the memory required grows only quadratically with k , whereas other algorithms presented in the literature (Herrero, 1994) have a cubic growth rate.

Another important feature of the algorithm is that a 'hashing' look-up technique was used to determine whether a newly generated node – a candidate for the next shell – was already in the middle or previous shell. This increased the speed of the program by about three orders of magnitude.

Optimizations with respect to both memory and speed were necessary: without them, it would not have been possible to compute the several hundred to several thousand CS terms that were required for the analysis. Computing times were a matter of hours on a high-speed workstation.

4. Algebraic description I: quadratic equations defining the topological density

It has already been shown in the literature (Brunner, 1979; Herrero, 1994; Schumacher, 1994; O'Keeffe, 1991a) that in many cases the k th term of the coordination sequence, N_k , increases quadratically with k (just as the surface of a sphere increases quadratically with its radius). A very simple example is SOD, where $N_k = 2k^2 + 2$ holds for all k . Brunner (1979) gives a more typical example: for diamond, 'the equation $N_k = 2.5k^2 + 1.75$ renders values too high by 0.25 if k is odd and too low by 0.25 if k is even'. An exact description for all k is achieved by introducing a periodic set of quadratic equations:

$$N_k = (5/2)k^2 + 3/2 \quad \text{for } k = 2n + 1, \\ n = 0, 1, 2, \dots$$

$$N_k = (5/2)k^2 + 2 \quad \text{for } k = 2n + 2, \\ n = 0, 1, 2, \dots$$

For ABW, the equations involve both k^2 and k :

$$N_k = (19/9)k^2 + (1/9)k + 16/9 \quad \text{for } k = 3n + 1, \\ n = 0, 1, 2, \dots$$

$$N_k = (19/9)k^2 - (1/9)k + 16/9 \quad \text{for } k = 3n + 2, \\ n = 0, 1, 2, \dots$$

$$N_k = (19/9)k^2 - 0k + 2 \quad \text{for } k = 3n + 3, \\ n = 0, 1, 2, \dots$$

In general, all coordination sequences investigated in this study can be described exactly by a periodic set of quadratic equations of the form

$$N_k = a_i k^2 + b_i k + c_i \quad \text{for } k = Mn + i, \quad (1) \\ \text{for } n = 0, 1, 2, \dots \quad \text{and} \quad i = 1, 2, \dots, M.$$

The number of equations, M , will be referred to as the period length. These equations hold for all k greater than or equal to some starting point k_0 .

Following a definition of O'Keeffe (1991a), we define the 'exact topological density' (TD) to be the mean of the a_i divided by the dimension N_D of the crystal space (which is 3 for all results presented here):

$$\text{TD} = \langle a_i \rangle / N_D = (1/MN_D) \sum_{i=1}^M a_i. \quad (2)$$

As k increases, the effect of the linear and constant coefficients b_i and c_i on the CS decreases and the quadratic coefficient a_i dominates. The error in the approximation

$$N_k \simeq N_D \text{TD} k^2 \quad (3)$$

vanishes for $k \rightarrow \infty$.

5. Algebraic description II: generating functions

The generating function (GF) for a coordination sequence (cf. Sloane & Plouffe, 1995),

$$\text{GF} = \sum_{k=0}^{\infty} N_k x^k, \quad (4)$$

often provides a more concise description than the quadratic equations given in (1). The generating functions for the sequences considered in this study have the form

$$\text{GF} = \sum_{i=0}^{O(IT)-1} IT(i)x_i \prod_{i=0}^{O(IT)-1} (1 - x^{PL(i)}), \quad (5)$$

where IT is a set of order $O(IT)$ 'initial terms' and PL is a set of (not necessarily distinct) $O(PL)$ 'period lengths'. The coefficients of the Taylor-series expansion of the generating function then give the CS. For example, the GF for ABW is

$$\text{GF}(\text{ABW}) = (1 + 3x + 6x^2 + 9x^3 + 9x^4 + 6x^5 + 3x^6 \\ + x^7) / [(1 - x^3)(1 - x^3)(1 - x)] \\ = 1 + 4x + 10x^2 + 21x^3 + 36x^4 \\ + 54x^5 + \dots$$

First, the GF's for some simple zeolites were obtained by using the Maple *GFUN* package (Waterloo Maple Software, 1994; Salvy & Zimmermann, 1994). However, more complex cases were beyond the capabilities of *GFUN* and an alternative approach was used. Note that the Taylor-series expansion of a generating function of the form (5) can be obtained by the simple recursive reconstruction algorithm shown in Fig. 1. This produces the coordination sequence from the sets IT and PL .

Conversely, given the set PL and a sufficient number of CS terms, the set IT can be obtained by multiplying the GF by the denominator of (5). This is accomplished by the recursive decomposition algorithm shown in Fig. 2.

6. Properties (P) and manipulations (M) of generating functions and connection with quadratic equations

(P1) The individual members of the set PL can be arranged in any order. This follows immediately from (5). On the other hand, the order of the elements of the set IT is important.

(P2) Extra period lengths can be adjoined to the set PL at the cost of enlarging the set IT . This corresponds to multiplying both the numerator and denominator of (5) by factors $1 - x^n$.

(P3) In some cases, it is possible to reduce the set PL by cancellation of factors. APD T1 gives a simple example: with $PL = \{21, 11, 8\}$, the set IT has 43 elements. However, with $PL = \{11, 8, 7, 3\}$, the set IT has 32 elements. No further simplification is possible.

(M1) Reduction of set IT with enlargement of set PL : Starting with initial sets IT/PL , two 'compact' sets IT_c/PL_c are obtained with the algorithm shown in Fig. 3, which is based on properties (P2) and (P3). This algorithm results in a set PL_c , of which no element can be omitted or further factorized. (However, this algorithm does not always produce the 'most compact' GF.)

```
Copy initial terms to N0...NO(IT)-1
Set NO(IT)...Nkmax = 0
For each i = 0 to O(PL)-1
  For each k = PL(i) to kmax
    Nk = Nk + Nk-PL(i)
```

Fig. 1. Recursive reconstruction algorithm.

```
Copy the CS terms to N0...Nkmax
For each i = 0...O(PL)-1
  For each k = kmax-1...PL(i) step -1
    Nk = Nk - Nk-PL(i)
```

Fig. 2. Recursive decomposition algorithm.

For IT_c/PL_c , the following relations hold for all CS's investigated:

(P4) Let IT_c and PL_c be the sets obtained with manipulation technique (M1). The starting point for the periodic set of quadratic equations can be taken to be the difference between the degrees of the numerator and denominator of (5) or, in other words,

$$k_0 = O(IT_c) - \sum_{i=0}^{O(PL_c)-1} PL_c(i). \quad (6)$$

(P5) Let IT_c and PL_c be the sets obtained with manipulation technique (M1). The least common multiple (LCM) of the elements of the set PL_c equals the period length of the set of quadratic equations:

$$M = \text{LCM}(PL_c).$$

For example, for EUO T9, we have

$$\begin{aligned} O(IT_c) &= 222 \\ PL_c &= \{36, 26, 24, 23, 22, 21, 17, 14, 10, 8, 5\} \\ \Rightarrow k_0 &= 222 - 206 = 16 \\ M &= 140\,900\,760. \end{aligned}$$

(M2) Reversal of (M1). Reduction of set PL with enlargement of set IT : The number of elements of an initial set PL can be reduced with the algorithm shown

in Fig. 4. Using this algorithm, it was possible to modify the generating functions for all the CS's investigated so that in every case exactly three PL elements were required.

(M3) Finding upper bounds on the subperiod lengths of the coefficients a_i, b_i, c_i of the set of quadratic equations: Upper bounds on the subperiod lengths of the coefficients a_i, b_i, c_i can be established with a very simple algorithm. Based on IT_c/PL_c , about $3\text{LCM}(PL_c)$ CS terms are computed. Next, the recursive decomposition algorithm is applied with $PL = \{\text{LCM}(PL_c), \text{LCM}(PL_c)\}$. With a linear period search in the resulting sequence, an upper bound pl_a on the subperiod length $O(a_i)$ is obtained. The process is repeated with $PL = \{\text{LCM}(PL_c), pl_a\}$ to obtain pl_b , and finally with $PL = \{pl_a, pl_b\}$ to obtain pl_c .

For all sequences except those of GOO T3, JBW T1 and T2, the Fe position of FeS_2 -marcasite, and both positions of $\text{CaF}_2(2)$, the bounds pl_a, pl_b and pl_c are exactly equal to the subperiod lengths $O(a_i), O(b_i)$ and $O(c_i)$, respectively. For the exceptional cases (and of course also in general), the relations $pl_a \geq O(a_i), pl_b \geq O(b_i)$ and $pl_c \geq O(c_i)$ hold. Moreover, for all CS's, the relation $pl_a \leq pl_b \leq pl_c$ holds, and $O(a_i) \leq O(b_i) \leq O(c_i)$ is valid for all CSs except those of GOO T3, JBW T1 and T2, the Fe position of FeS_2 -marcasite, and both positions of $\text{CaF}_2(2)$.

By application of the manipulation techniques (M2) and (M3), it was always possible to obtain $O(PL) = 3$ for

```
(a) Test if any element of PL can be omitted, by seeing if the
    corresponding term 1-x^n divides the numerator of (5).
    If so, remove the element.
(b) Loop through the elements of PL.
    Call the current element the pivot element.
    Loop through all integer divisors of the pivot element
    Copy PL and divide the pivot element by the divisor
    Compute a large number of CS terms from the initial IT/PL
    Apply the recursive decomposition algorithm and do a
    linear period search to recover a potentially missing
    last period length
    Upon successful period search, restart with step (a)
```

Fig. 3. Reduction of set IT with enlargement of set PL .

```
(a) Test if any element of PL can be omitted.
    If so, remove the element.
(b) Loop through all pairs of elements of PL
    Copy PL, omitting the current pair of elements
    Compute a large number of CS terms from the initial IT/PL
    Apply the RD algorithm and do a linear period search to
    recover a potentially missing last period length
    Upon successful period search, restart with step (b)
(c) Repeat (a) and (b) until no change occurs
```

Fig. 4. Reduction of set PL with enlargement of set IT .

all 390 three-dimensional CS's that were investigated. It is conjectured that this holds for any three-dimensional CS.

In simple cases, such as the f.c.c. or b.c.c. lattices, the elements of the set IT are positive integers. Indeed, it follows from the work of Stanley (1976, 1980) that if certain conditions are satisfied (one of which is that the set of points in or on the k th coordination shell is convex), then the IT are necessarily positive. In the present investigation, however, many examples with negative IT elements were encountered. For example, the As position of NiAs(1) has $IT = \{1, 4, 12, 10, -5, 2\}$, $PL = \{2, 1, 1\}$.

6.1. Application of (M2) and (M3)

Since the computation of the CS terms with the recursive reconstruction algorithm requires the whole sequence to be held in memory, major difficulties arise for large period lengths. For example, the determination of the topological density for EUO requires about $3M = 422\,702\,280$ 64-bit integers to be stored, a total of about 3.15 Gbytes. However, for the case $O(PL) = 3$, a special-purpose algorithm for computing the CS terms with small memory was devised. This algorithm is several orders of magnitude slower than the simpler recursive reconstruction algorithm, but – in combination with the manipulation techniques (M2) and (M3) – enabled the determination of the topological density even for the largest cases.

7. Second differences of CS's

The second derivative of a quadratic polynomial is a constant. By analogy, the second differences between successive terms of the CS in a number of examples are constant, or have a constant period. For example, the CS of the SiO_2 polymorph tridymite (tri) and the corresponding first ($F_k = N_k - N_{k-1}$) and second ($S_k = F_k - F_{k-1}$) differences are

$$\begin{aligned} \text{CS}(\text{tri}) &= \{4, 12, 25, 44, 67, 96, 130, 170, 214, 264, \\ &\quad 319, 380, 445, 516, \dots\} \\ \text{diff}_1[\text{CS}(\text{tri})] &= \{8, 13, 19, 23, 29, 34, 40, 44, 50, 55, 61, \\ &\quad 65, 71, \dots\} \\ \text{diff}_2[\text{CS}(\text{tri})] &= \{(5, 6, 4, 6), (5, 6, 4, 6), (5, 6, 4, 6), \dots\}. \end{aligned}$$

The period length of the second differences is four, as indicated by the parentheses. [Although it is not needed here, it is worth mentioning that Herschel's 'circulator' notation provides a convenient terminology for describing such periodic sequences (Comtet, 1974, p. 109).]

The occurrence of a constant period in the second differences implies $b_i = 0$ for the entire (periodic) set of

quadratic equations. For example, for ABW, some b_i are different from zero (see above) and the numerical values of the second differences are not constant. However, their plot (Fig. 5) clearly reveals a periodicity that can be used to estimate one period length for the recursive decomposition. In this case, the period length is 3. And, indeed, the set of period lengths for ABW is $\{3, 3, 1\}$.

8. Strategy for the determination of the exact topological density

Typically, the determination of the exact topological density (TD) requires the following steps:

- (i) Computation from 100 to 2000 terms of the CS with the node-counting algorithm.
- (ii) Investigation of the second differences; often this reveals one 'period length' for the recursive decomposition.
- (iii) Determination of the period lengths for the recursive decomposition.

(iv) Alternative 1: Computation of a few million terms of the CS, search for a periodic set of quadratic equations, and computation of TD from (2). Alternative 2: Computation of k_0 via (6), $M = \text{LCM}(PL_c)$ and the corresponding $3O(a_i)$ CS terms; computation of one subperiod of a_i ; computation of TD from (2).

All but step (iii) have been outlined above. For the determination of the period lengths for the recursive decomposition, two techniques have been applied:

(a) A set of n maximum period lengths $PL_{\max} = \{pl(0)_{\max}, pl(1)_{\max}, \dots, pl(n-1)_{\max}\}$ is prescribed and the first position is tested from 1 to $pl(0)_{\max}$. At each pass of the loop, the recursive decomposition algorithm is applied to the CS, followed by a check for a sufficiently large sequence of zeros at the end of the resulting sequence. Also, a linear period-seeking algorithm tries to recover a possibly last missing period length. In the

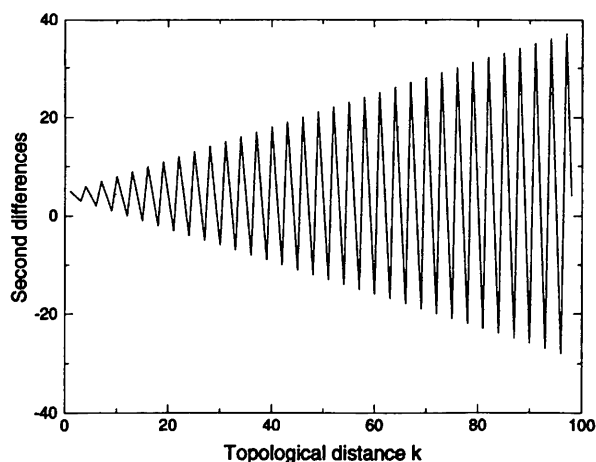


Fig. 5. Second differences of ABW.

case of no success, $pl(1)$ is also tested and a loop with two variables, but avoiding permutations, is executed. Unless a solution has been found, more and more loop positions are activated until the entire set is depleted.

(b) A review of several known PL sets revealed that individual pl often occur in pairs. This observation and the exploitation of properties ($P1$) and ($P2$) led to the following strategy: given a large number of CS terms, attempt to obtain a decomposition using $PL = \{m, m, m - 1, m - 1, \dots, 2, 2, 1, 1\}$.

Upon success, the initial set PL is subjected to manipulation technique ($M1$), in order to obtain IT_c and PL_c , the compact form of description of the particular CS.

8.1. Example with decomposition technique (a)

For EUO T9, 1460 CS terms were calculated with the node-counting algorithm. Investigation of the second differences suggested a pl of 15, which was then applied to the CS. The resulting sequence was processed with $PL_{\max} = \{400, 300, 200, 100, 40, 40, 40\}$ (of course, there was no hope that this loop would ever run to completion). At loop position $PL = \{391, 286, 40\}$, the linear period-seeking algorithm found a period with length 504. Using the combined set $PL = \{504, 391, 286, 40, 15\}$, the size of the set IT is 1252. By means of the manipulation technique ($M1$), the final solution $PL_c = \{36, 26, 24, 23, 22, 21, 17, 14, 10, 8, 5\}$ was obtained, together with 222 IT_c elements.

8.2. Example with decomposition technique (b)

For MAZ T2, 999 CS terms were calculated with the node counting algorithm. Setting $PL = \{15, 15, 14, 14, \dots, 1, 1\}$ revealed a solution with 241 IT elements. By means of manipulation technique ($M1$), the final solution $PL_c = \{14, 11, 10, 8, 7\}$ with 51 IT_c elements was obtained.

9. Results of the algebraic analysis

For 127 crystal structures, 402 coordination sequences were investigated, of which 390 are unique. For the zeolite structures, there are eight sequences that occur in two or more crystallographically different environments: ABW = ATN; LTA = RHO; CAN = AFG T2 = AFG T3 = LIO T2 = LIO T4 = LOS T2; GME = AFX T1; AFS T3 = BPH T3; EDI T1 = THO T1; ERI T1 = OFF T1; EAB T2 = OFF T2. Furthermore, the CS of one atom of α -Nd (the position at the origin) is equal to the CS of Mg, and one CS of NiAs(1) (again for the position at the origin) is equal to the CS of W(1).

Tables 3–5 list the results of the algebraic analysis. Column S gives the number of different CS's for the structure indicated in the first column. Except for AFG and LIO, which have two different sites with the same

CS, this is also the number of crystallographically distinct positions. $O(IT)$ and $O(PL)$ designate the order of the set of initial terms and period lengths, respectively. For structures with more than one topologically distinct site, the range is given (e.g. EUO: 213–233 initial terms). However, if there are only two values, these are separated by a comma instead of a hyphen [e.g., for EUO, $O(PL)$ is either 10 or 11 for all ten sequences], and in some cases all sets have an equal number of members and only one value is necessary. M is the number of quadratic equations that make up the periodic set. TD, the exact topological density defined by (2), is given exactly, as a rational fraction multiplied by $N_D = 3$ ($TDN_D = \langle a_i \rangle$) and, for better comparability, also as decimal number. TD10 and Δ (%) are defined by (7) and (8) below.

All the coordination sequences in this study have been added to the electronically accessible version of Sloane & Plouffe (1995) at the address sequences@research.att.com (Sloane, 1994). In this way, the CS can be used as a 'fingerprint' to assist in the identification of a crystal structure.

10. Properties of the coefficients of the quadratic equations

For most of the zeolites with only one or two tetrahedral sites, and for the dense SiO_2 polymorphs, the structure of the coefficients of the quadratic equations was investigated. These structures all share one or more of the following characteristics: the parameter a is the same for all terms or shows a shorter period than b and c ; the parameter b is zero or shows a shorter period than c ; different periods appear for odd and even values of i ; some or all parameters are 'palindromic'. The following are some typical examples. The most 'special' example is SOD. The CS is $N_k = 2k^2 + 2$: 4, 10, 18, 34, ..., thus, $a = 2$, $b = 0$, $c = 2$; the period M is 1. This type of equation also occurs for the structures of NaCl, W and Cu (P , I and F lattice complexes) with $a = 4$, 6 and 10, respectively.

A somewhat less special example is AFI. The parameters are $a = 21/10$, $b = 0$ for all values of k and $M = 10$. The parameter c is palindromic about $k = M/2 = 5$. This means that c is the same for $k = 1$ and $k = 9$; it is also the same for the pairs 2 and 8, 3 and 7, and 4 and 6, respectively. For $k = 10$, the parameter c is 2. Such palindromic behavior about $k = M/2$ is frequently observed. An example in which M is odd is KFI with $a = 12/7$, $b = 0$ for all k , $M = 7$, where c is palindromic about $k = 3.5$ (thus, c is the same for the pairs $k = 1$ and 6, 2 and 5, and 3 and 4, respectively). For the last term of the period ($k = 7$), c is again 2. Another kind of palindromic behavior may appear if b is not zero: the magnitudes of b and of c are the same for pairs of k as in the previous examples but the signs of b may differ for the two members of a pair. An example is CAN.

Table 3. Results for topologies listed in the Atlas of Zeolite Structure Types (1992, 1996)

Code	S	O(IT)	O(PL)	M	TD frac. $\times 3 = \langle a_i \rangle$	TD dec.	TD10	Δ (%)
ABW	1	8	3	3	19/9	0.703704	833.0	2.36
AEI	3	36, 45	4	1320, 2640	2309/1320	0.583081	688.7	2.11
AEL	3	29, 33	3	36, 72	497/216	0.766975	903.8	1.91
AET	5	47-72	3, 4	168, 336	67/32	0.697917	824.1	2.11
AFG	2(3)	12, 27	3	5, 20	52/25	0.693333	815.5	1.71
AFI	1	13	3	10	21/10	0.7	828.0	2.29
AFO	4	35, 49	3, 4	48	665/288	0.769676	907.4	1.96
AFR	4	62, 85	5, 6	31395, 125580	163664/94185	0.579229	686.7	2.49
AFS	3	39, 52	4, 5	120, 240	273/160	0.56875	655.7	0.34
AFT	3	31	3	420	123/70	0.585714	684.7	1.06
AFX	2	17, 24	3	140	123/70	0.585714	688.5	1.63
AFY	2	19, 20	3, 4	60	22/15	0.488889	585.2	3.46
AHT	2	11, 23	3	8	35/16	0.729167	853.3	1.20
ANA	1	17	3	40	12/5	0.8	933.0	0.87
APC	2	16, 31	3	45, 360	94/45	0.696296	814.0	1.09
APD	2	22, 32	3, 4	231, 1848	526/231	0.759019	887.5	1.12
AST	2	16	4	12	15/8	0.625	742.2	2.68
ATN	1	8	3	3	19/9	0.703704	833.0	2.36
ATO	1	12	3	5	57/25	0.76	894.0	1.73
ATS	3	14-19	3	15, 30	48/25	0.64	752.3	1.64
ATT	2	23	4	420	68/35	0.647619	767.7	2.50
ATV	2	20	3	40	49/20	0.816667	960.3	1.70
AWW	2	18, 29	3	63, 504	124/63	0.656085	772.3	1.78
*BEA	9	236-257	11, 12	742560	81978419/38785500	0.704545	805.1	1.19
BIK	2	19, 20	5	12	49/18	0.907407	1052.3	0.31
BOG	6	81-94	7	16720	113787/57475	0.659922	780.8	2.31
BPH	3	39, 44	4, 5	120	273/160	0.56875	667.3	1.43
BRE	4	64, 67	6	2340	12289/5265	0.778031	900.5	0.10
CAN	1	12	3	5	52/25	0.693333	817.0	1.90
CAS	3	33, 40	5, 6	120	7741/2880	0.895949	1042.3	0.63
CHA	1	11	3	20	17/10	0.566667	677.0	3.28
-CHI	4	103-112	8	198968, 397936	153225069/61282144	0.833441	913.3	5.23
-CLO	5	62, 68	4	18480	512/385	0.44329	455.5	11.23
CON	7	96-101	7, 8	102960	3105307/1544400	0.670229	784.0	1.15
DAC	4	62-65	10	9240	4896737/1940400	0.84119	977.3	0.49
DDR	7	120-129	9	55440	39307/15400	0.850801	967.9	1.61
DFO	6	79-100	4	53010	15268/8835	0.576042	663.6	0.41
DOH	4	85-94	6	120120	145707/55055	0.882191	1001.9	1.77
EAB	2	19, 25	3	210, 420	66/35	0.628571	735.0	1.10
EDI	2	9, 11	3	12	2	0.666667	786.2	1.97
EMT	4	66, 75	4, 5	3360	2071/1400	0.493095	584.0	2.37
EPI	3	44	8	420	89441/35280	0.845059	978.7	0.17
ERI	2	19, 25	3	210, 420	66/35	0.628571	738.3	1.56
EUO	10	213-233	10, 11	140900760	395365279/150965100	0.872973	964.9	4.40
FAU	1	15	3	42	10/7	0.47619	579.0	5.09
FER	4	47-62	6, 7	420, 840	55921/21000	0.887635	1021.4	0.47
GIS	1	9	3	12	11/6	0.611111	726.0	2.72
GME	1	17	3	140	123/70	0.585714	694.0	2.44
GOO	5	31-40	5	420	2032/945	0.716755	840.2	1.37
HEU	5	28-40	5	420	4903/2100	0.778254	908.6	0.97
JBW	2	19	4	20	113/50	0.753333	890.3	2.21
KFI	1	10	3	7	12/7	0.571429	681.0	3.03
LAU	3	26, 28	4	1260	622/315	0.658201	782.0	2.73
LEV	2	24	4	210	318/175	0.605714	719.0	2.63
LIO	3(4)	12, 22	3	5, 15	52/25	0.693333	815.7	1.74
LOS	2	12, 17	3	5, 10	52/25	0.693333	816.0	1.77
LOV	3	45-54	7	660	34233/15125	0.754446	879.2	0.78
LTA	1	8	3	5	8/5	0.533333	641.0	3.90
LTL	2	25	3	504	13/7	0.619048	746.0	4.20
LTN	4	139-177	5	251940, 503880	3384/1615	0.698452	779.2	3.53
MAZ	2	51, 53	5	3080	11271/5390	0.697032	823.0	2.10
MEI	4	110, 121	7, 8	5419260	301573/159390	0.630682	727.9	0.21
MEL	7	74-80	7	80080	121417/50050	0.808638	944.1	0.98
MEP	3	88, 91	7	3527160	421222/146965	0.955379	1058.8	4.14
MER	1	10	3	15	28/15	0.622222	738.0	2.55
MFI	12	185-235	7, 8	62622560	96965483/39139100	0.825819	959.9	0.53

Table 3 (cont.)

Code	<i>S</i>	<i>O(IT)</i>	<i>O(PL)</i>	<i>M</i>	TD frac. $\times 3 = \langle a_i \rangle$	TD dec.	TD10	Δ (%)
MFS	8	51-76	7, 8	420, 840	127349/49000	0.86632	994.8	0.68
MON	1	26	5	12	187/108	0.885802	1033.0	0.87
MOR	4	93-95	8	32760	298988/124215	0.80234	938.3	1.14
MTN	3	58, 60	6	1560	4522/1625	0.92759	1049.1	2.17
MTT	7	118-135	9	13860	17672791/6670125	0.883181	1015.0	0.60
MTW	7	77-86	8, 9	13860	3194357/1372140	0.776004	911.7	1.61
NAT	2	27, 29	3, 4	24	20/9	0.740741	834.2	2.61
NES	7	104-124	7	159390	352325/143451	0.818688	922.1	2.59
NON	5	85-98	8	32760, 65520	321277/117000	0.915319	1037.5	1.96
OFF	2	19	3	210	66/35	0.628571	739.0	1.65
-PAR	4	49-58	7, 8	6930, 13860	518384/259875	0.664915	773.2	0.55
PAU	8	29-44	3	77, 154	144/77	0.623377	728.1	0.99
PHI	2	23, 28	4	60	143/75	0.635556	750.5	2.10
RHO	1	8	3	5	8/5	0.533333	641.0	3.90
-ROG†	3	46, 55	5, 6	240	1063/600	0.590556	690.0	1.01
-RON	4	54, 63	6	240	7441/3600	0.688981	771.1	3.23
RSN	5	83, 90	9	8580	5570407/2359500	0.786947	913.6	0.40
RTE	3	20, 23	4	420	451/210	0.715873	844.3	1.99
RTH	4	53, 61	7, 8	8190, 16380	384632/184275	0.695757	816.7	1.51
RUT	5	74-80	8, 9	4680	1293083/561600	0.767499	902.1	1.65
SGT	4	68-81	6, 7	120	13979/5400	0.862901	962.2	3.56
SOD	1	4	3	1	2	0.666667	791.0	2.60
STI	4	40-43	5	1560	2107/975	0.720342	851.9	2.27
THO	3	9-15	3	12, 24	2	0.666667	784.2	1.71
TON	4	38-51	6	60, 120	1951/750	0.867111	1005.7	0.32
VET	5	83-92	8, 9	3276, 6552	544799/198744	0.913737	1023.4	3.12
VFI	2	24, 37	3	56, 112	27/16	0.5625	668.7	2.77
VNI	7	254-283	11, 12	16432416	611375421655/227293178112	0.896603	971.0	6.33
VSV	3	48, 55	6	60	1227/500	0.818	948.1	0.24
WEI	2	20	4	12	425/216	0.655864	773.4	1.96
-WEN	3	29, 34	3, 4	770, 2310	148/77	0.640693	755.0	1.89
YUG	2	22, 25	4, 5	105, 420	754/315	0.797884	935.0	1.35
ZON	4	48, 56	4	9660	328/161	0.679089	797.7	1.57

† The structure-type code -ROG has been discredited (see IZA Structure Commission Report, 1994) and is included only for completeness.

Table 4. Results for dense SiO_2 polymorphs

Mineral	<i>S</i>	<i>O(IT)</i>	<i>O(PL)</i>	<i>M</i>	TD frac. $\times 3 = \langle a_i \rangle$	TD dec.	TD10	Δ (%)
Banalsite	2	12	3	12	17/6	0.944444	1053.0	3.56
Coesite	2	36, 40	6	420	1363/378	1.20194	1318.5	5.10
Cordierite	2	26, 35	5	60, 120	279/100	0.93	1058.3	1.57
Cristobalite	1	5	3	2	5/2	0.833333	981.0	1.82
Feldspar	2	8, 10	3	3	20/9	0.740741	874.0	2.04
Keatite	2	34, 35	4	120	81/25	1.08	1225.7	1.82
Milarite	2	22, 27	3	120	21/8	0.875	1004.2	0.73
Moganite	2	8, 10	3	3	22/9	0.814815	958.3	1.72
Paracelsian	1	8	3	5	11/5	0.733333	864.0	1.89
Quartz	1	9	3	6	19/6	1.05556	1231.0	0.89
Scapolite	2	31, 33	5	252	941/378	0.829806	975.0	1.62
Tridymite	1	7	3	4	21/8	0.875	1027.0	1.52

Very often, the parameters *a* and *b* are the same for all values of *k* or show a much shorter period than *c*. An example is AFY: for site T1, *b* = 0 for all *k*, for atom T2, *b* has period 3 with the values 1/6, -1/6, 0, respectively, while *c* has a period of 30 if *k* is odd and 60 if *k* is even. In the examples mentioned so far, either *b* itself or the sum of the *b*'s over the period is zero.

This is not the case for either atom of milarite: all values of *b* are negative.

10.1. TD as topological invariant

In those cases where the parameter *a* is not the same for all values of *k* but is periodic, $\text{TD} = \langle a_i \rangle / N_D [(2)]$ is the same for all atoms in a framework. A check with two-dimensional nets showed that (after a certain number of spheres) it is even the same if any cluster is chosen as a 'starting point'. Therefore, for all structures considered in this work, the CS's have been computed

Table 5. Results for selected non-tetrahedral structures

Formula	S	O(IT)	O(PL)	M	TD frac. $\times 3 = (a_i)$	TD dec.	TD10	Δ (%)
CaF ₂ (2)	2	7	3	2	15/4	1.25	1487.7	2.97
CaF ₂ (1)	2	8, 10	3	3, 6	80/9	2.96296	3410.3	0.38
NaCl	1	4	3	1	4	1.33333	1561.0	1.30
FeS ₂ -Mar	2	10, 16	4	6, 12	47/12	1.30556	1523.7	0.98
FeS ₂ -Pyr	2	7, 13	3	2, 4	9/2	1.5	1716.3	0.99
NiAs (2)	2	5, 7	3	2, 4	9/2	1.5	1748.0	0.84
NiAs (1)	2	4, 6	3	1, 2	6	2	2325.0	0.61
Cu	1	4	3	1	10	3.33333	3871.0	0.52
Mg	1	5	3	2	21/2	3.5	4061.0	0.43
W (2)	1	4	3	1	12	4	4641.0	0.43
W (1)	1	4	3	1	6	2	2331.0	0.87
α -Nd	2	5, 7	3	2, 4	21/2	3.5	4058.0	0.36
Ni ₂ In	2	7	3	6	11	3.66667	4251.0	0.35
α -Mn	4	95, 102	6	17160	1562201/102960	5.05763	5354.5	8.36
Cr ₃ Si	2	15, 17	3, 4	12	187/12	5.19444	5579.5	7.02
σ -CrFe	5	162-184	11	6846840	1829724773/112972860	5.39871	5609.5	10.06
MgZn ₂	3	38, 52	5, 6	315, 630	3978/245	5.41224	5704.3	8.76
MgCu ₂	2	18, 19	6	12	2371/144	5.48843	5788.3	8.71
MgNi ₂	5	61-83	7-9	210, 840	123787/7350	5.61392	5800.7	10.55

and analyzed with all sites in the unit cell as starting cluster. In any case, the resulting TD was equal to the TD when starting with only one atom. Thus, we conclude that TD is actually a topological invariant of a framework {in general, the parameters a , b , c and M [(1) are not}, comparable to the cycle classes introduced by Beukemann & Klee (1994).

10.2. Correlation of TD and TD10

The *Atlas of Zeolite Structure Types* (1992, 1996) lists a quantity called the 'topological density' denoted by TD10. This is defined to be the average number of nodes (atom sites) in a cluster of topological radius 10, weighted by the multiplicities m_j of the s sites:

$$\text{TD10} = \left[\sum_{j=1}^s m_j \left(1 + \sum_{k=1}^{10} N_{kj} \right) \right] / \sum_{j=1}^s m_j. \quad (7)$$

Use of (3) (TD is the same for all sites of a structure) leads to

$$\begin{aligned} \text{TD10} &\simeq \left[\sum_{j=1}^s m_j \left(1 + \sum_{k=1}^{10} (\text{TD} N_D) k^2 \right) \right] / \sum_{j=1}^s m_j \\ &= 1 + (\text{TD} N_D) \sum_{k=1}^{10} k^2 \\ &= 1 + 385(\text{TD} N_D) \end{aligned}$$

and hence

$$\text{TD} \simeq (\text{TD10} - 1) / 385 N_D = \text{TD10}_{\text{norm}}. \quad (8)$$

The rightmost columns of Tables 3-5 [Δ (%)] list the percentage deviations of $\text{TD10}_{\text{norm}}$ from the exact TD. Fig. 6 is a histogram of the distribution of the deviations. For the majority of the structures, the deviation is well below 3% but there are also some outliers. A deviation

of more than 5% was found for -CHI and -CLO, two interrupted frameworks, for FAU, and for six of the non-tetrahedral structures. These three zeolites represent relatively open and/or complex structures and one might conclude that more than ten steps are necessary in such cases in order to achieve a satisfactory convergence, but on the other hand very complex structures like PAU and MFI show a very good correlation between the exact topological density and TD10. In view of this, the additional effort necessary to obtain the exact TD seems justified, and previous work based on approximations to the exact value should perhaps be reconsidered.

10.3. Consequences for the definition of the coordination sphere

The choice of the bond length (Table 2) determines the network and is a constant matter of debate. The

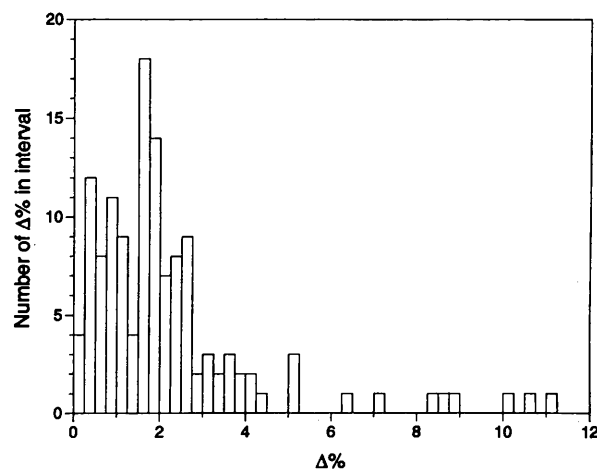


Fig. 6. Histogram of TD/TD10 correlation (see Tables 3-5).

quantities TD and TD10, which reflect long-range properties of the framework, may help in choosing a sensible specification for the local environment. As an example, the 8-coordinated network of tungsten [W(1) in Tables 2 and 5] leads to strange values for the densities, while the 14-connected network has densities that fit in the list of metallic structures. The 14-connected network admits the second-longest bonds, which are only 15% longer. A similar geometric situation exists in CaF_2 but now the second-nearest (also 15% more distant) neighbors of fluorine should not be admitted. They would lead to values for the densities that differ considerably from those of metals, whereas the lower coordination with only four neighbors brings CaF_2 (2) in the vicinity of NaCl, and again related structures stand together.

11. Some special n -dimensional periodic graphs

The coordination sequence of an n -dimensional graph or net (cf. Wells, 1977) necessarily has $O(PL) \geq n$.

In one dimension, the simplest periodic graph is a linear chain, for which the CS is $1, 2, 2, \dots$, with $IT = \{1, 1\}$, $PL = \{1\}$. In two dimensions, the hexagonal net 6^3 (which occurs for example in the mineral biotite) has the CS $1, 3, 6, 9, 12, 15, \dots$, with $IT = \{1, 1, 1\}$, $PL = \{1, 1\}$. In three dimensions, as already mentioned, sodalite has $IT = \{1, 1, 1, 1\}$, $PL = \{1, 1, 1\}$. In a sense, these are the simplest coordination sequences in dimensions 1, 2 and 3.

O'Keeffe (1991b) has generalized these three structures to higher dimensions, defining ' n -dimensional sodalites' for all n . He also gives their coordination sequences for $n \leq 6$. The generating functions of these sequences have been analyzed and were found to continue the pattern of the first three: the set IT consists of $n + 1$ 1's and PL of n 1's. It is reasonable to conjecture that this holds in general. In a forthcoming paper (Conway & Sloane, 1996), it will be shown that this conjecture is equivalent to the assertion that the points in or on the k th coordination shell of n -dimensional sodalite are in one-to-one correspondence with the n -dimensional 'centered tetrahedral' numbers.

12. Conclusions

The main objectives of this work were to obtain (i) an exact definition and numerical values for topological densities, which can be used to investigate correlations with other properties of the crystal structures, and (ii) a better understanding of coordination sequences, since these are now frequently used to characterize structures [see, for example, the novel structure-determination technique in Grosse-Kunstleve (1996)]. Regarding this aim, we consider our work to be fully successful. We have developed a recursive decomposition for the fast

and efficient calculation of an arbitrary number of CS terms and have computed exact numerical values for the topological density, which is an invariant of all the structures investigated.

However, the results are empirical, as there is no rigorous mathematical proof that a generating function of the form (5) must hold for the CS of a periodic structure. The applicability of (5) has been verified for certain one-, two- and three-dimensional periodic topologies. In the case of one and two dimensions, the justification is straightforward but already in three dimensions there are difficulties. This is a relatively minor point. Once the generating function has been discovered, watching its Taylor-series expansion match the CS for hundreds or even thousands of terms carries complete conviction!

Another unresolved question concerns the conditions under which the special numerical properties of the quadratic equations hold.

The help of Uwe Hollerbach at Boston University in running the period search for the coordination sequences of EUO and MFI is very much appreciated.

References

- Akporiaye, D. E. & Price, G. D. (1989). *Zeolites*, **9**, 321–328.
Atlas of Zeolite Structure Types (1992). 3rd ed., edited by W. M. Meier & D. H. Olson. London: Butterworth-Heinemann.
Atlas of Zeolite Structure Types (1996). 4th ed., edited by W. M. Meier, D. H. Olson & C. Baerlocher. Amsterdam: Elsevier.
 Barthomeuf, D. (1993). *J. Phys. Chem.* **97**, 10092–10096.
 Beukemann, A. & Klee, W. E. (1994). *Z. Kristallogr.* **209**, 709–713.
 Brunner, G. O. (1979). *J. Solid State Chem.* **29**, 41–45.
 Brunner, G. O. & Laves, F. (1971). *Wiss. Z. Tech. Univ. Dresden*, **20**, 387–390.
 Comtet, L. (1974). *Advanced Combinatorics*. Dordrecht: Reidel.
 Conway, J. H. & Sloane, N. J. A. (1995). *Discret. Comput. Geom.* **13**, 383–403.
 Conway, J. H. & Sloane, N. J. A. (1996). *Proc. R. Soc. London Ser. A*. In the press.
 Fischer, W. (1973). *Z. Kristallogr.* **138**, 129–146.
 Fischer, W. (1974). *Z. Kristallogr.* **140**, 50–74.
 Grosse-Kunstleve, R. W. (1996). PhD Dissertation, ETH Zurich, Switzerland.
 Herrero, C. P. (1993). *Chem. Phys. Lett.* **215**, 587–590.
 Herrero, C. P. (1994). *J. Chem. Soc. Faraday Trans.* **90**, 2597–2599.
 Inorganic Crystal Structure Database (1986–1995). Universität Bonn, Germany.
 IZA Structure Commission Report (1994). *Zeolites*, **14**, 389–392.
 Meier, W. M. & Möck, H. J. (1979). *J. Solid State Chem.* **27**, 349–355.
 O'Keeffe, M. (1991a). *Z. Kristallogr.* **196**, 21–37.
 O'Keeffe, M. (1991b). *Acta Cryst.* **A47**, 748–753.
 Salvy, B. & Zimmermann, P. (1994). *ACM Trans. Math. Software*, **20**, 163–177.

- Schumacher, S. (1994). Dissertation, Universität Karlsruhe, Germany.
- Sloane, N. J. A. (1994). *Electron. J. Comb.* **1**, No. 1.
- Sloane, N. J. A. & Plouffe, S. (1995). *The Encyclopedia of Integer Sequences*. New York: Academic Press.
- Stanley, R. P. (1976). *Duke Math. J.* **43**, 511–531.
- Stanley, R. P. (1980). *Ann. Discret. Math.* **6**, 333–342.
- TYPIX (1994). *Gmelin Handbook of Inorganic and Organometallic Chemistry*, 8th ed. Berlin: Springer-Verlag.
- Waterloo Maple Software (1994). *Maple V*, Release 3, a language for symbolic mathematical calculation. University of Waterloo, Canada.
- Wells, A. F. (1977). *Three-Dimensional Nets and Polyhedra*. New York: Academic Press.